Dialectologia 16 *(2016), 143-157.*
ISSN: 2013-2247

Received 22 April 2014.
Accepted 12 June 2014.

LINGUISTIC MAPS & DIALECT DATA PROCESSING[1*]

Deokho KIM[1] & Sanggyu LEE[2.]

Kyungpook National University, South of Korea.

idhkim@knu.ac.kr[1] / sglee@knu.ac.kr[2]

Abstract

The present study explores the major accomplishments made thus far with special consideration of the path of development linguistic map making has taken hitherto in Korea and Japan. The purpose of creating linguistic maps is to identify linguistic boundaries, either big or small, and assign systematic meanings as well as determine the relationship between spatial distribution and change by time. The linguistic maps should be understood in their broader concept that shows the various geographic information related to linguistic branches within a particular geographic area as well as relevant trends and aspects in linguistic branches on not only paper but also online screens and images. This study classifies linguistic maps into manual method maps and computerized maps, for which the time of computerization has been identified as a standard for classification, and explores their types focusing on those demonstrating high performance in presentation techniques.

Keywords

Linguistic map, Dialect map, computerized linguistic map, Audio linguistic map, Aialect data, Sejong Plan

MAPAS LINGÜÍSTICOS Y PROCESAMIENTO DE DATOS DIALECTALES

Resumen

El presente estudio explora los principales logros alcanzados hasta el momento, con especial atención en el desarrollo que los mapas lingüísticos han tenido hasta ahora en Corea y en Japón. El objetivo de la creación de mapas lingüísticos es identificar las fronteras lingüísticas, ya sean grandes o pequeñas, y asignar significados sistemáticos, así como determinar la relación entre la distribución

[1]* This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government(NRF-2013S1A5A8025734).
·1. First author 2. Corresponding author.
∗ Kyungpook National University Department of Korean Languae and Literature. 80 Daehakro, Bukgu, Daegu, 41566, South of Korea.

espacial y el cambio temporal. Los mapas lingüísticos deben ser entendidos en su concepto más amplio, ya que muestran, no sólo en papel, sino también en pantallas y en imágenes en línea, información geográfica diversa relacionada con distintas ramas de la lingüística en una área geográfica particular, y también las tendencias y aspectos lingüísticos relevantes. Este estudio clasifica los mapas lingüísticos en mapas de métodos manuales y mapas informatizados, para los que el tiempo de la informatización se ha identificado como un estándar para la clasificación, y explora sus tipologías centrándose en los que demuestran un alto rendimiento en las técnicas de presentación.

Palabras clave

Mapa lingüístico, mapa dialectal, mapa lingüístico informatizado, audio mapa lingüístico, datos dialectales, Plan de Sejong

1. Introduction

The purpose of creating linguistic maps is to identify linguistic boundaries, either big or small, and assign systematic meanings as well as determine the relationship between spatial distribution and change by time. Hence, linguistic maps gather the linguistic data collected from a certain research area and "present the basic material in the map" (Macaulay 1985: 175).

According to the traditional notion of a linguistic map, "a linguistic map is to express the linguistic differences within a particular region in the form of a map" (The Society of Dialectology 2001: 265). However, the linguistic maps should not be understood in this traditional notion that shows the linguistic branches in the form of a map, but they should be understood in their broader concept that shows the linguistic characteristics on three-dimensional online screens as well as on the maps or schematics.

2. Types of linguistic maps

Linguistic maps published around the world or distributed online have come in various forms. In general, linguistic maps are categorized as i) display maps and

interpretive maps depending on their form of presenting data; and ii) qualitative maps and quantitative maps depending on their way of producing maps.

The types of linguistic maps have been further diversified since computers have gained popularity. Linguistic maps can be more exhaustively classified into a variety of types in accordance with the quantity and form of data, the way of presenting data, the level of comprehensiveness and simplification of data processing, the form of visual presentation of data, etc. Meanwhile, the differences among diverse types of linguistic maps are explained by means of the different presentation techniques used depending on the performance levels of computers as well as the advancements made in software of databases or graphic systems, or the different developmental stages in creating linguistic maps and linguistic data. The present study classifies manual method maps and computerized maps, and examines their types focusing on those which demonstrate high performance in presentation techniques.

2.1. Manually created linguistic maps

Manually created linguistic maps are produced by displaying dialect transcripted form or symbolic markers manually in each area on a blank map. The manually created linguistic maps are further classified into two types; i) offset method through which the manually created linguistic maps are printed offset (the same method is applied to master printing) and ii) tracing method through which a researcher manually draws a map and a printing designer transfers them to a tracing paper.

Such manual methods of creating linguistic maps are uneconomical and also lack the credibility in accuracy. Those linguistic maps produced by mimeograph machines around 1950s are not only crude but also barely legible.

2.2. Computerized linguistic maps

Map-making methods using PCs, which organize dialect data and transfer them to maps, are very effective in enhancing the analytical capacity of linguistic data.

The types of linguistic maps are diversified according to the size of the research area, the scale of a map, the number of research areas, the types of data, the methods of map-making, etc. Furthermore, the methods of data processing and the ways of computer-based map-making are also diverse. Computer-based map-making methods either utilize the previously developed software programs or develop and use a fully computerized map-making system.

2.2.1. Linguistic maps using the existing software

The process of computerization of linguistic map-making in Korea can be condensed into a search for dialect data collecting methods, techniques of creating images on a blank map and printing them out, symbolization of dialect data and displaying them on a blank map, etc. At the beginning, such word processors as 'OHP', 'boseokgeul', and 'hanguel' were mainly used in computerizing the data. Mail merge or the method through which grids were established and the areas were specified on a blank map was mainly employed for image display and print-out. Recently previously developed software programs (bitmap, paint brush, photo shop, illustrator) are utilized for image-making on a blank map and print-out. Yet exhaustive software has not been developed which enables computers to cover the whole process from linguistic map-making to print-out in establishing data base for dialect data.

Jung (정인상) (1985) first used dBASE II in creating linguistic maps, while Kim, Hong, Kim & So (김충회,홍윤표,김병선,소강춘) (1991) used dBASE III. As Kim (김덕호) (2002)'s linguistic map making method using mail merge has been widely distributed, the method of making isogloss maps using graphic software programs like Photoshop, ArcView, mass flow controller for windows, etc. has spread out.

2.2.2. Grid-matrix maps

A location map is based on a detailed the map of basis, while in a grid-matrix map the overall structure of the regions or the map of basis are abstractly represented and systematized in the grid. Kim (김충회) (1992) introduced the Grid-matrix technique of

an exclusive linguistic map making system using Boseokgeul II and HANDB into the Korean academic world [Figure 1].

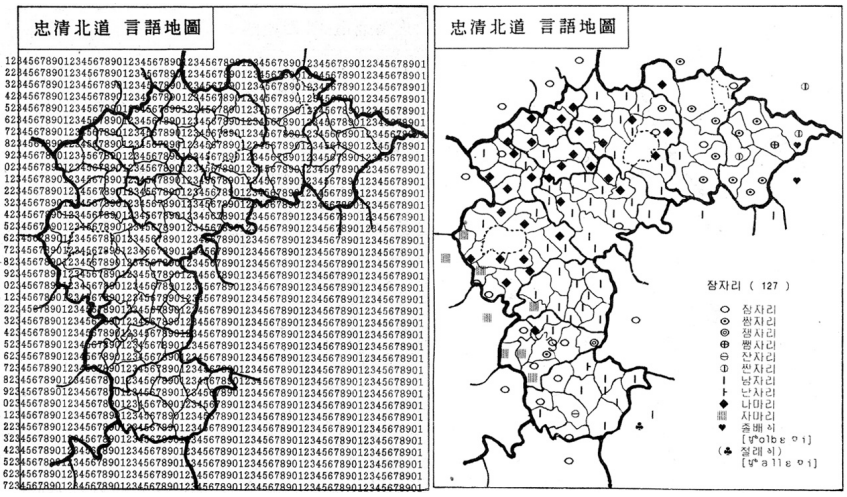

Figure 1. Kim (김충회) (1992)'s linguistic map of ChungCheongbuk-do

In Japan, Ogino Tsunao (荻野綱男) developed a linguistic map-making software called GLAPS (Ogino 1994). GLAPS produces maps which are expressed in cross-table method and glottogram method. In this program, 80 by 80 columns of grid are encoded and converted into diskette files. Takahashi-genzi (高橋顯志)'s 'Boring linguistic map' arrays graph paper-like grid patterns on a blank map and draws the linguistic distribution on each area.

Actually this is a two dimensional expressive method. However it not only expresses the information of many different strata but also suggests a frequency glottogram technique which displays information of the informant. SYMVU (designed by Harvard University's Computer Graphic Laboratory), a British linguistic map-making program, is another location of map-making method, which was utilized for the Welsh dialect data analysis by Alan R. Thomas (1980). Likewise a cartogram which is a type of grid map and cluster-type maps has been widely used for data processing of special information since GIS program was released.

2.2.3. Dialect search and linguistic maps

A 'dialect search program' was developed as a part of Sejong Plan (2001), which not only searches dialect data but also identifies linguistic distributions on a map [Figure 2]. The program was developed by two professional researchers, that is Tu & An (두길수,안동언) (2002). And this Plan was supported by National Institute of the Korean Language in the South Korea.

This program is very useful for searching for dialect data. It marks dialect distributions on a linguistic map in colour, employing the unique Arcview's mask technique. The linguistic Map produced by the search system has a couple of drawbacks; i) it falls behind in accuracy, and ii) it cannot input the data. However, the graphic search is possible.

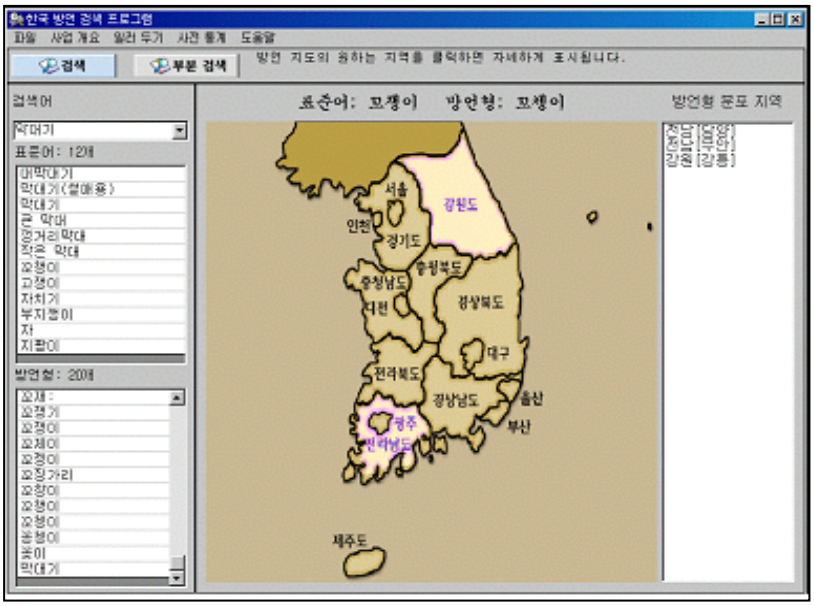

Figure 2. Korean dialect search system by Sejong Plan (2001)

2.2.4. Symbolic sign linguistic maps

Symbolic sign linguistic maps use various symbolic signs and colours on a map. Among others, the National Institute for Japanese Language started to computerize the whole process of making symbolic sign linguistic maps from 'Grammar Atlas of Japanese Dialects (GAJ, 方言文法全國地圖)' vol. 5. A GAJ is created by folding many layers of

image of maps and letters, signs and colours of dialects. Since a GAJ basically uses Adobe Illustrator, but signs in each area are made to automatically select a form of differentiation in dialect, it can be considered to be a fully automatic program for linguistic map making.

Another type of symbolic sign linguistic map is SEAL system developed by Fukushima (1983) [Figure 3]. This system also operates in Windows English versions. It can process the linguistic data from other countries and create maps for them. Furthermore it can create interpretative maps which reflect the frequencies of lexicon. Lee (이상규) (2003) developed KSEAL by adding three more letters to SEAL in order to transcribe dialect data in Korean [Figure 4]. Symbolic sign linguistic maps are the most efficient in expressing geological distribution of linguistic data, but it has a limit as to reflecting detailed areal information causing differentiation in language. And Lee (이상규) (2005) developed MAP MAKER (a linguistic Map making tool), which can make symbolic sign linguistic maps in a easier way in Windows [Figure 5].

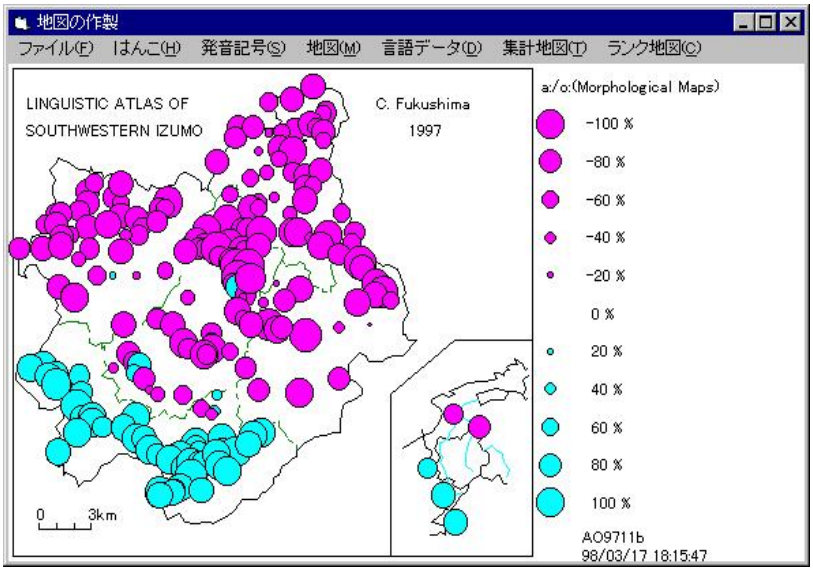

Figure 3. Fukushima Chitsuko (1983)'s SEAL

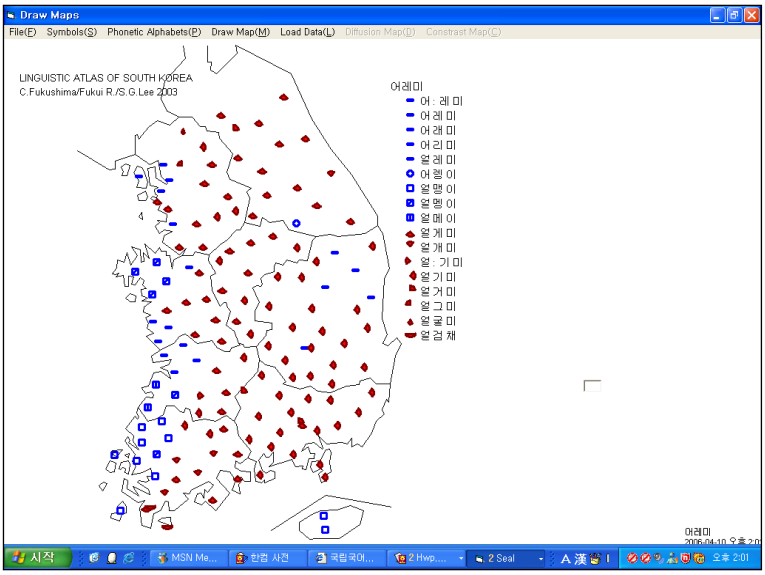

Figure 4. Lee (이상규) (2003)'s KSEAL

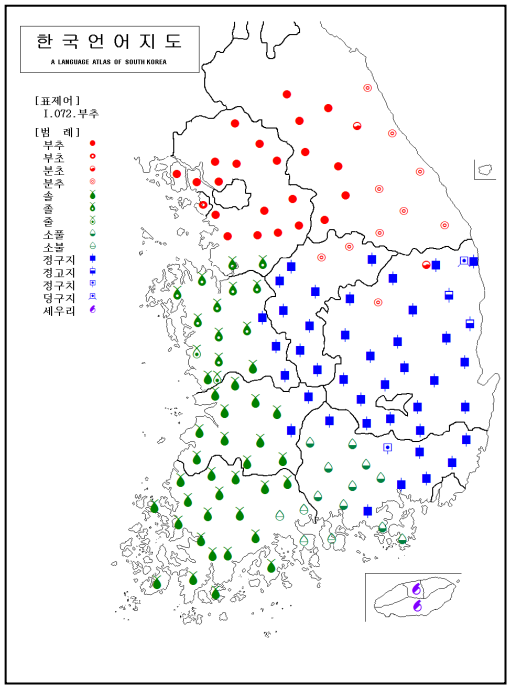

Figure 5. Lee (이상규) (2005)'s MAP MAKER

2.2.5. Audio linguistic maps

Audio linguistic map is made by simultaneously displaying both encoded linguistic forms and voices of the informants collected at the spots on a linguistic map. It uses previously developed software and adds voice information to the linguistic map.

Dialectologia 16 *(2016), 143-157.*
ISSN: 2013-2247

Such linguistic maps are created by using previously developed software like Excel 2000, File maker Pro 5.5, Acrobat 5.0, etc. in the procedure of ① date input, ② making sign fonts, ③ converting into signs, ④ making a blank map, ⑤ making new file using File maker Pro, ⑥ making PDF format and programming of fonts. And then Acrobat link property is utilized to hyperlink each single sign by the link tool and connect them with the voice file.

Audio linguistic map was developed by Kishie Shinsuke (岸江信介) at University of Tochushima as a linguistic map making program for natural discourse data [Figure 6]. Nakai Seiichi (中井精一) at University of Toyama developed an Audio linguistic map making system using GIS.

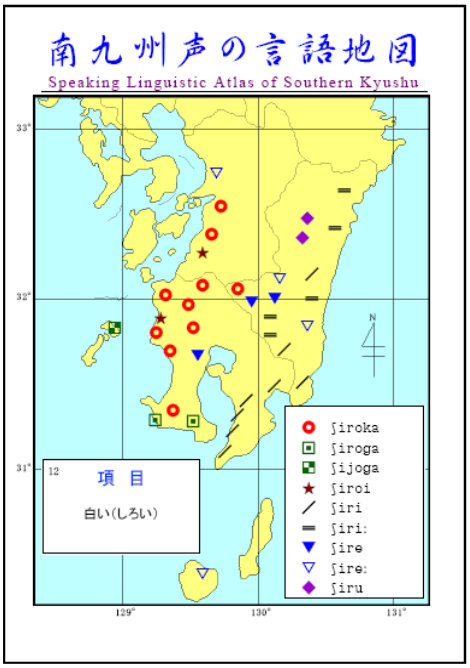

Figure 6. Kishie Shinsuke (岸江信介)'s Audio linguistic map

In the South of Korea, N. I. K. L. (국립국어원) (2007) further plans to develop a tool which can systematically manage computerized programs for voice data as well as data of dialect transcription including Audio linguistic map production (making) based on MAP MAKER [Figure 7].

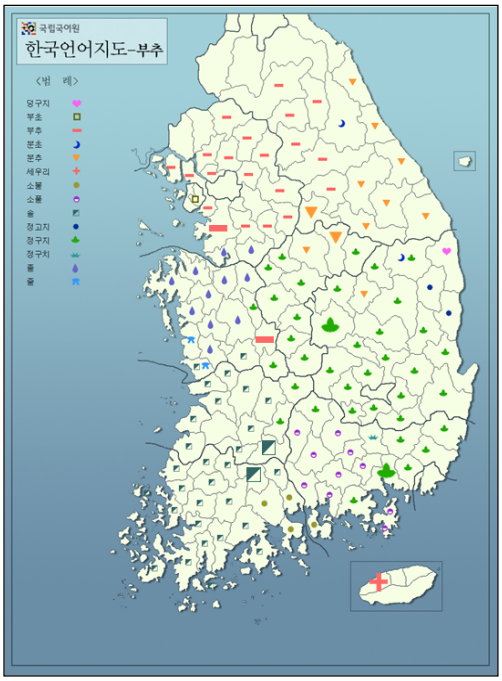

Figure 7. N. I. K. L. (2007)'s Audio linguistic map

2.2.6. Dialect perception maps

A perception map is a type of linguistic Map which is attempted to divide the dialects by examining the linguistic perception of the dwellers or reflect the dialectal differences on a map based on socio-perceptual differences. Dennis R. Preston (1989: 25) surveyed the perception level of the American dialect among people living in Hawaii and divided by their language and dialects. Perceptual Dialect Maps of Korea were developed by Lim (任榮哲) (1992) with Daniel Long and Isao Onishi. Kim Deokho (2001) and Kim, Kishie & Takiguchi (2012) integrated the native speaker's cognitions in the Gyeongsangbuk-do of the South Korea and made the Perceptual dialect map of non-linguistic features [Figure 8].

Dialectologia 16 *(2016), 143-157.*
ISSN: 2013-2247

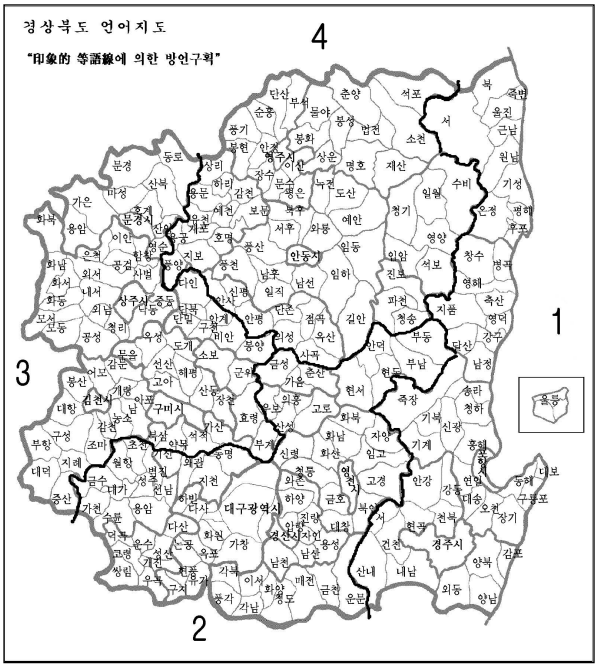

Figure 8. Kim (2001) and Kim, Kishie & Takuguchi (2012)'s the Perceptual dialect map of Kyungpook Korea

2.2.7. Multi-dimensional linguistic maps using GIS

The GIS-Research at University of Toyama founded <人文科學と GIS> in 2003. It has tried to interpret the factors for differentiation in language multi-dimensionally, using MANDARA, a supporting system for geological information analysis. Especially the ways to develop the techniques to use the geographic information system (GIS) in making linguistic maps as interdisciplinary studies among folklore, socio-linguistics, etc. will be actively examined. In the future Audio linguistic map making utilizing the geographic information system should be attempted in a full scale. Currently Nakai Seiichi (中井精一) and Kibe Nobuko (木部暢子) at University of Toyama are actively discussing the methods to make Audio linguistic map.

3. Tools for linguistic Map making

3.1. Development of tools for linguistic Map making

The methods of making linguistic maps are rapidly developing. It has developed from the direct input of dialect form on a map, via the symbolic map-making method by which dialect form is converted into symbolic signs (signs, numbers, colours, etc.) to the method by which the informants' voice data are simultaneously presented on a map. Furthermore three-dimensional geographic features are also used as the basis of linguistic maps, which show linguistic distributions on a map with such various geographical marks as rivers and mountains expressed on it. The presentation techniques in linguistic map making using geographic information system have reached the level of more than three dimensional world, synthesis of time and space, and presenting the virtual environment of linguistic geography by simulating the time and space of the past and future. In the near future linguistic maps will make it easy to search voice data and also make it possible to connect them with transcripted data.

Many analysis tools for the study of language have been introduced in the world, and recently various tools for linguistic maps using graphic tools and Excel have been introduced. I myself introduced SEAL in the linguistic world of Korea. However, the SEAL was originally developed in Japan, thus it has many limits.

3.2. Characteristics of linguistic Map maker

The linguistic Map Maker has been introduced in Kim (2010). This program is the tool for making symbolic sign linguistic maps and has many advantages as follows.

1) It makes it possible to produce linguistic maps very easily in Windows. When clicking 'new file creation' in the menu, a pop-up comes on, which leads to open the necessary files to make maps according to order so that one can easily produce maps just by following the procedure the program suggests. Furthermore this program has a setting function which minimizes the time of map-making. One can set for the use of basic blank maps, the use of basic coordinates, and the use of random symbolic signs by

clicking file -> setting change in order. Such basic setting of use is a very useful function in making many different word maps with the same blank maps. It prevents repeating the same procedure to make new maps and thus can minimize the time in making maps.

2) The analysis and interpretation of dialect data can be simplified through its arithmetic and sorting function. In other words, the program has the function which sorts dialect form in the alphabetical order or in the order of frequency and makes it possible to catch the types of dialect differentiated form and the frequency of occurrence at a glance, and thus helps to select which words should be put on dictionaries.

3) The automatic coordinate setting function makes it possible to produce the type of maps the user wishes. When a basic map is used for a map-making as a base, the desired basic map is scanned and coordinates are built one by one on the blank map, using image tools. However, the map maker's function to generate coordinator data automatically makes it easy to produce desired types of maps.

4. Concluding remarks - New strategy for base of dialect data

Until now linguistic maps have been produced based on the generation of data base on each area after local data are surveyed, and then interpretation of language differentiation has been attempted. However, it is necessary to seek for a new method for base of dialect data. First, word networks should be constructed by systematizing the data collected on the spot. Language differentiation should be interpreted based on these word networks, and then linguistic maps should be produced accordingly.

The construction of word networks should be done by composing the semantic relationships among the words using facet-system. Likewise word networks should be constructed according to 1) etymological factors of language differentiation, 2) morphological factors of language differentiation, and 3) phonetic and phonological factors of language differentiation based on the dialect data. The interpretation of language differentiation should be attempted at the stage of base of the data. Accordingly clearer linguistic maps can be created.

References

JUNG, I. S. (정인상) (1985) "컴퓨터를 이용한 방언자료의 처리방안", *방언 8*, 한국정신문화연구원.

KIM, C. H. (김충회) (1992) *충북방언의 언어지리학*, 인하대출판부, 173-225.

KIM, C. H., Y. P. HONG, P. S. KIM & K. C. SO (김충회,홍윤표,김병선,소강춘) (1991) "방언자료의 전산처리에 대한 연구", *한국어전산학(창간호)*, 한국어전산학회.

KIM, D. H. (김덕호) (2001) "인상적 등어선에 대한 연구", *어문학* 제 73 집, 한국어문학회, 1-20.

KIM, D. H. (김덕호) (2002) *경북방언의 지리언어학*, 월인.

LEE, S. G. (이상규) (2003) *컴퓨터를 활용한 언어지도 제작 시스템*, 경북대학교 언어지도 제작연구실.

LEE, S. G. (이상규) (2005) "방언 자료의 처리와 언어지도", *방언학(창간호)*. 한국방언학회.

LIM, Y. C. (任榮哲) (1992) *在日.在米韓國人および韓國人の言語生活の實態*, 日本くろしお出版.

N. I. K. L. (국립국어원) (2007) *Audio linguistic map,* The National Institute of The Korean Language.

THE SOCIETY OF DIALECTOLOGY (2001) *The Dictionary of Dialectology*, THEHAKSA.

TU, G. S. & D. E. AN (두길수,안동언) (2002) "한국 방언 검색 프로그램 개발", *한국어와 정보*, 태학사.

田原廣史 (2002) コンピータと方言研究, *21 世紀の方言研究*, 國書刊行會.

井上史雄 (2002) "言語地圖의 魅力", *日本語學* 9 月.

高橋顯志 (2002) "Boring 言語地圖", *日本語學* 9 月.

大西拓一郎 (2002) "言語地圖製作の電算化", *日本語學* 21.

富山大學人文學部 GIS 研究會 (2003) *人文科學と GIS*, 富山大學人文學部 GIS 研究會.

中央大學校日本學研究所 (2004) "東アゼア聲の言語地圖の作成に向けて", 聲の言語地圖研究發表會.

FUKUSHIMA, Chitsuko (1983) *Pasokon niyoru Gengo-chirigaku eno Aputoochi (An Approach to Linguistic Geography Assisted by a Personal Computer-SEAL User's Manual)*, Private edition.

KIM, Deokho (2010) "Production of a dialect map utilizing map maker", *Dialectologia.* Special issue, I. Universitat de Barcelona, 21-46.
http://www.publicacions.ub.edu/revistes/dialectologiaSP2010/

KIM, Deokho, Shinsuke KISHIE & Keiko TAKIGUCHI (2012) "The Study of Perceptual dialectology in Kyungpok dialects", *Journal of Language and Literature Volume XX*, The Faculty of Integrated Arts and Sciences, The University of Tokushima, 117-137.

MACAULAY, R. K. S. (1985) "Linguistics Map: Visual Aid or Abstract Art?", in J. M. Kirk, S. F. Sanderson & J. D. A. Widdowson (eds.), *Studies in linguistic Geography*, Croom Helm, 172-186.

OGINO, Tsunao (荻野綱男) (1994) *GLAPS no Tsukaikatk (How to use GLAPS)*, Private edition.

PRESTON, Dennis R. (1989) *Perceptual Dialectology: Nonlinguists' Views of Areal Linguistics*, Dordrecht, the Netherlands: Foris Publications Holland.

THOMAS, Alan R. (1980) *Areal Analysis of Dialect Data by Computer*, Cardiff: University of Wales Press.