

Is it possible to measure the distance between near languages ? - A case study of French and Japanese -



Yuji KAWAGUCHI

Tokyo University of Foreign Studies

Introduction

What is the “nearness” for languages ?

Japanese	English
hoogen (方言)	dialect
namari (訛り)	accent

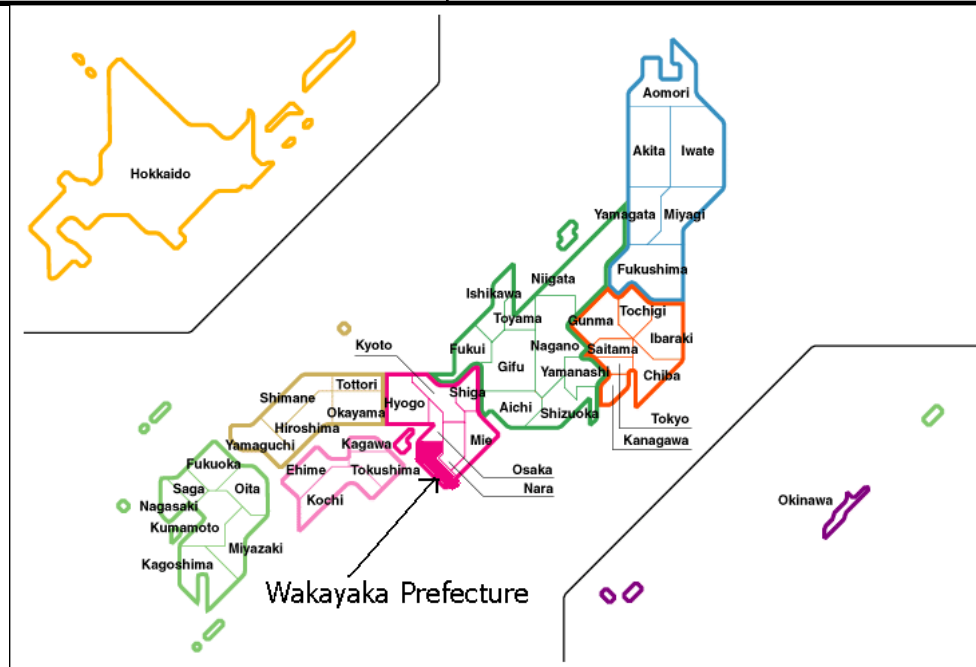
Furusato-no *namari* natsukashi teishaba-no
hitogomi-no naka-ni so-wo kiki-ni yuku.

“Railway station filled with lovely local accents,
I go into the crowd and listen to them.”

(Poem of Takuboku ISHIKAWA, 1886-1912)

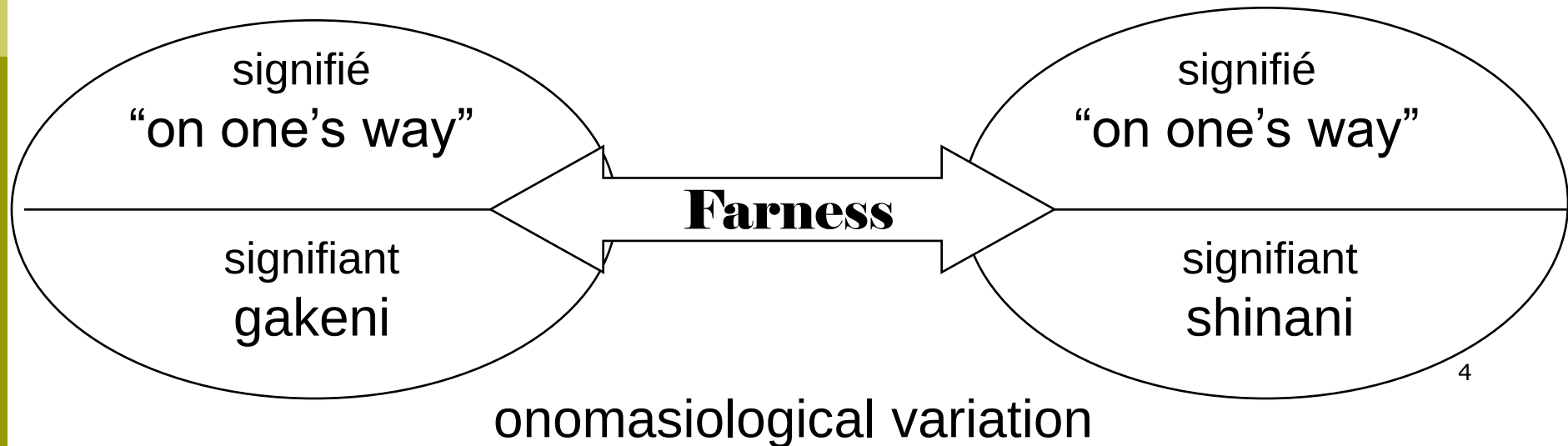
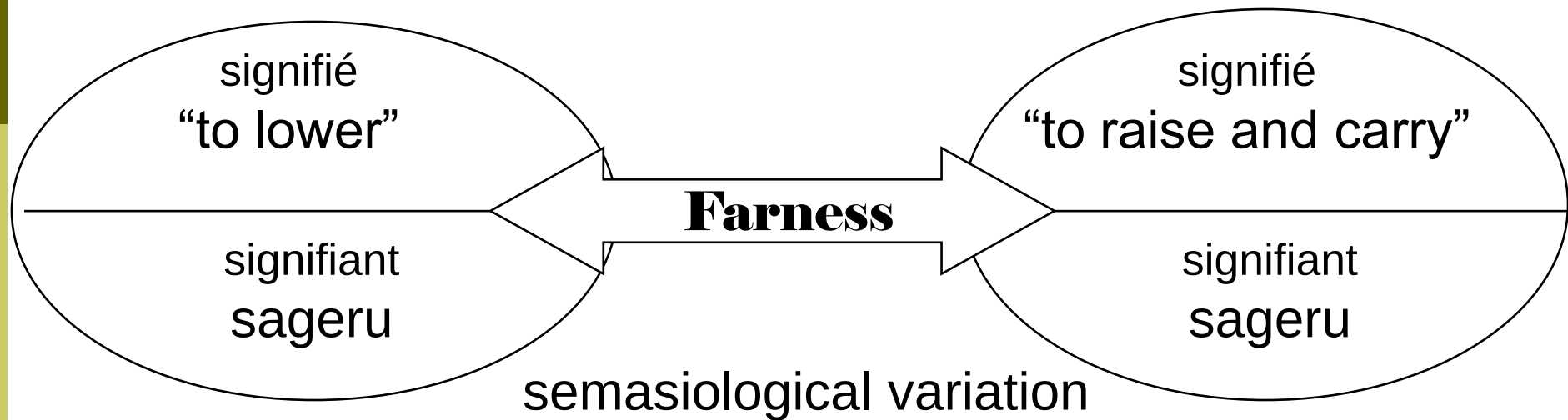
1. What is the “nearness” or “farness” in linguistic variation ?

standard Japanese	Wakayama variant
sageru “to lower” gakeni “on one’s way” mi-nai “not to look at” soshite “and then”	sageru “to raise and carry” shinani “on one’s way” mi-yan “not to look at” hoshite ~ hoite “and then”



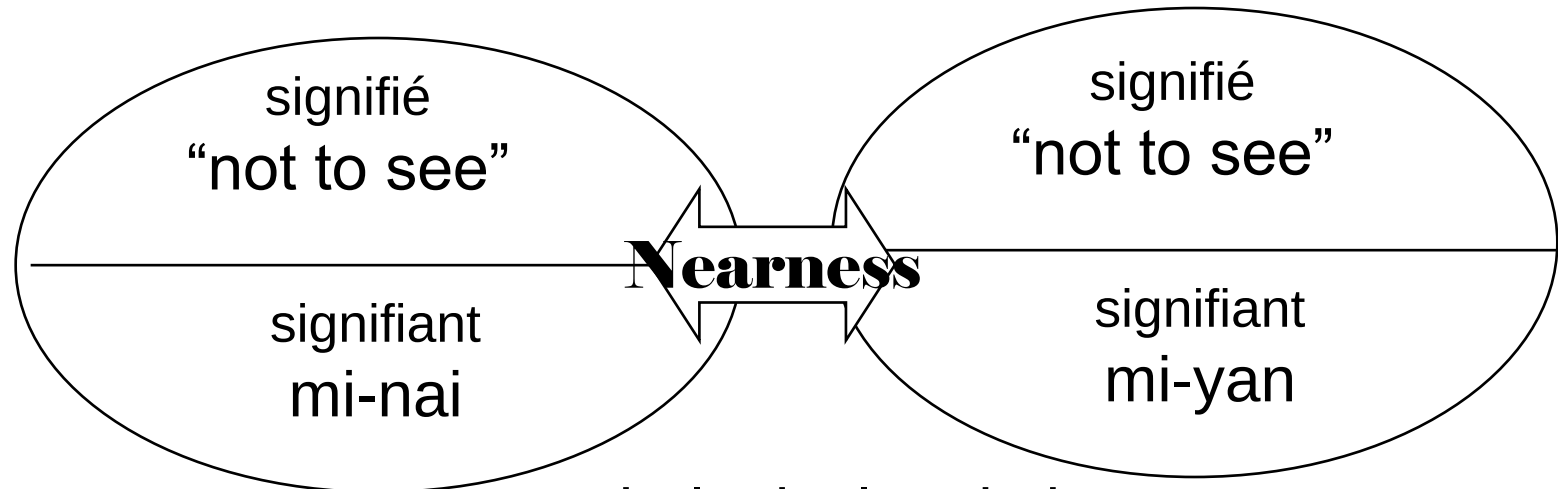
Type of variation I

Lexical variation

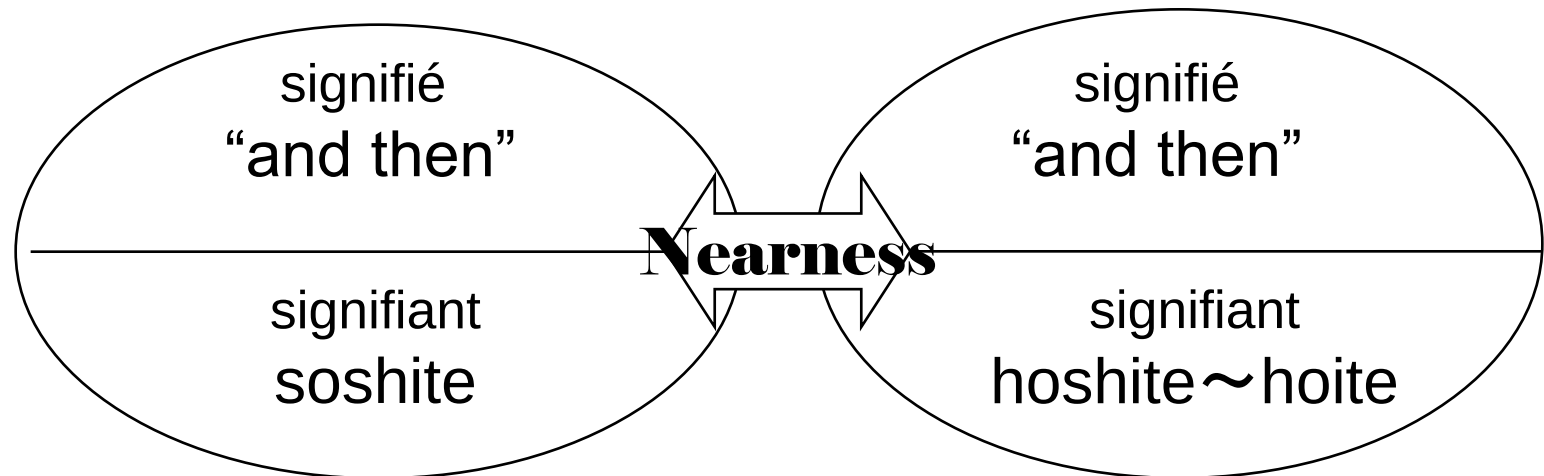


Type of variation 2

Morpho-phonological variation



morphological variation



phonological variation

2. How can we measure linguistically the distance of geographical variants ?

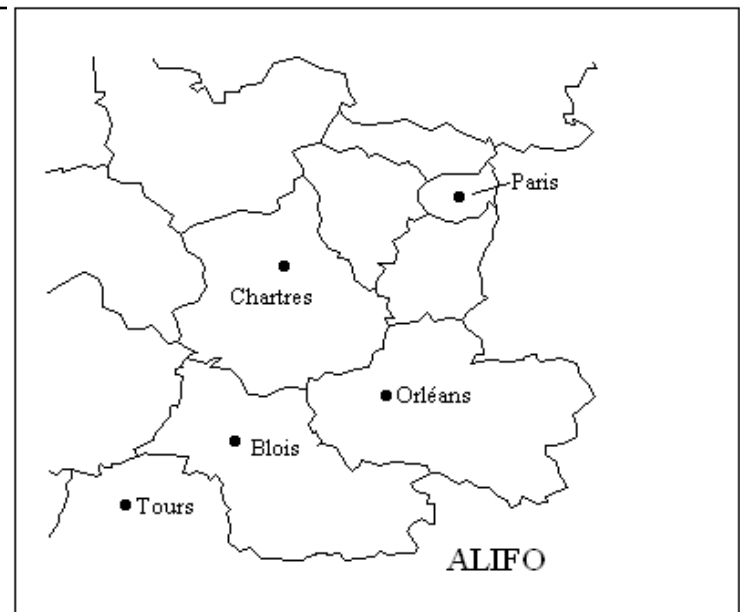
No.62 faner	“to make hay”
fáné, fàné	standard form
fànè, fèné, fènè, fèné	phonological variants
dèròtè, ègáyè	lexical variants

No.256 cerise	“cherry”
sèriz	standard form
sèriz	phonological variants
g’iñ, gin, giñ	lexical variants

No.440 vent du nord	“north wind”
vã-dü-nòr	standard form
vã-dü-nór, vã-dü-nó	phonological variants
vã-dü-ó, vã-ó	morphological variants

ALIFO (*Atlas linguistique et ethnographique de l’Île-de-France et de l’Orléanais*), Marie-Rose Simoni-Aurembou : 1966, vol.1, 1969, vol.2, 1978, vol.3.

Standard form is based on the description in Martinet, A. and H. Walter, 1973: *Dictionnaire de la prononciation française dans son usage réel*, France-Expansion.



Cluster Analysis for measuring distance between geographical variants

Purpose

1. to show if the distance between geographical variants can be measured through experimental multivariate analysis
2. to find the linguistic cue for nearness or farness in the distinction between lexical and morpho-phonological variation

Cluster analysis

Evaluation of geographical variants

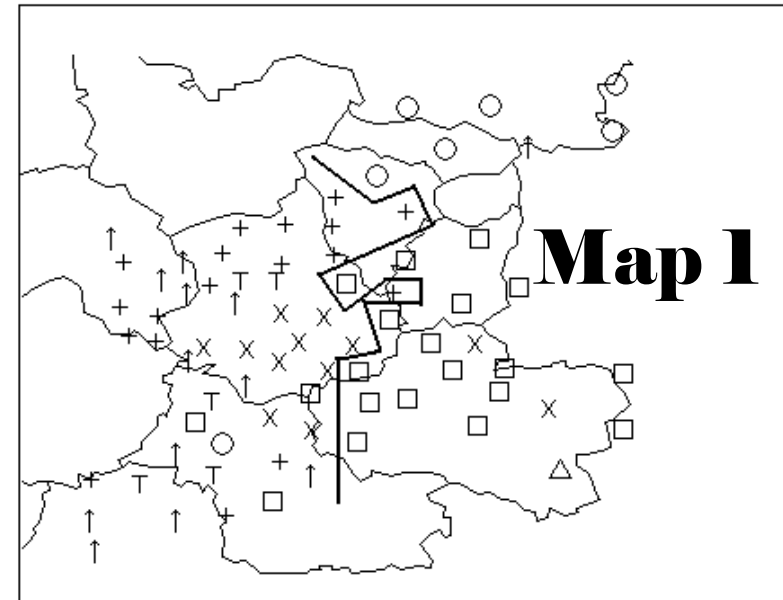
- Case 1: Nominal scale 1 or 2
value 1 for standard form, 2 for the others
- Case 2: Ordinal scale 1-4 or 5
value 1 for standard, 2-4 for ordering variants, in some cases 5
- Case 3: Ordinal scale 1-10 with a special weighting
value 1 for standard, 2-5 for ordering variants
10 exclusively for lexical variants

3. How will multivariate analysis illustrate the distance ?

No.62 charrue		Case 1	Case 2	Case 3	
	category	value	value	value	occurrences
Çàrü	standard	1	1	1	6
Çárü	phonological variants	2	2	2	59
Çá ^e rü		2	2	3	1
Çárü-vèrswèr		2	3	4	2
Çárü _y		2	3	4	1
Çèrü _y		2	3	3	1
Ça ^o rü		2	2	3	6
Çèrü		2	3	3	2
bràkér	lexical variants	2	4	10	1
kutèt		2	4	10	1
vèrswè		2	4	10	1
vèrswa ^e r		2	4	10	1

Area of standard form

CASE 1
Complete Linkage Method
Manhattan Distance

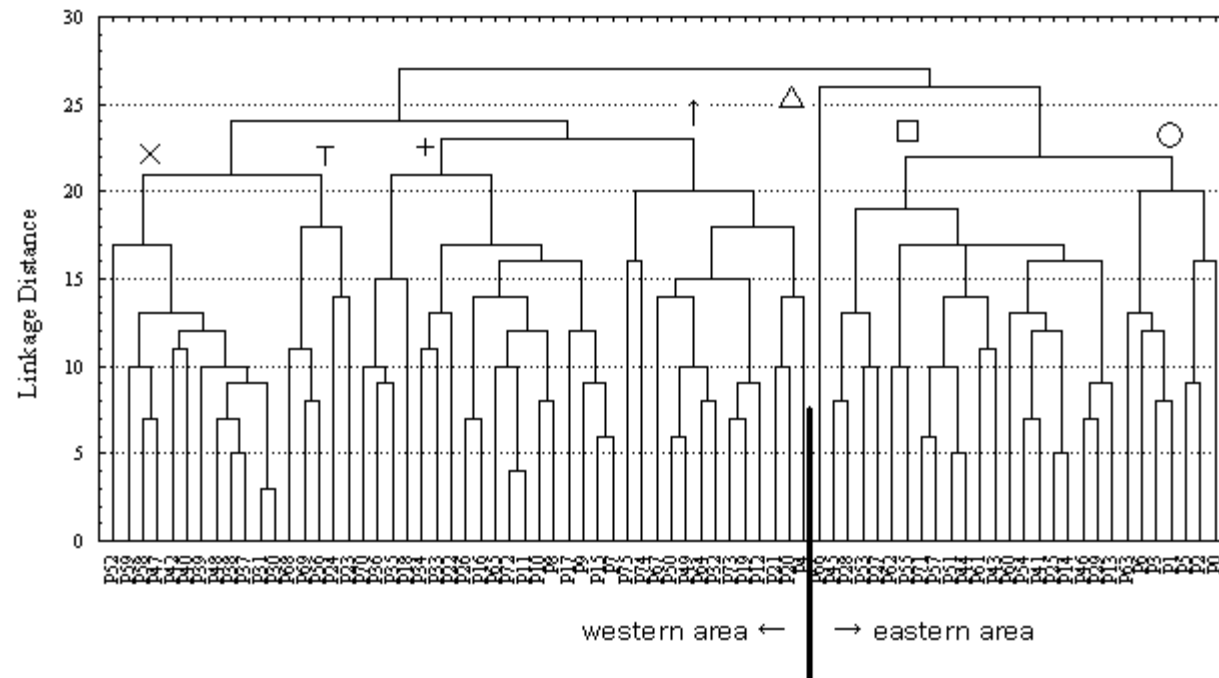


Nominal scale 1 or 2

Cluster O □ △

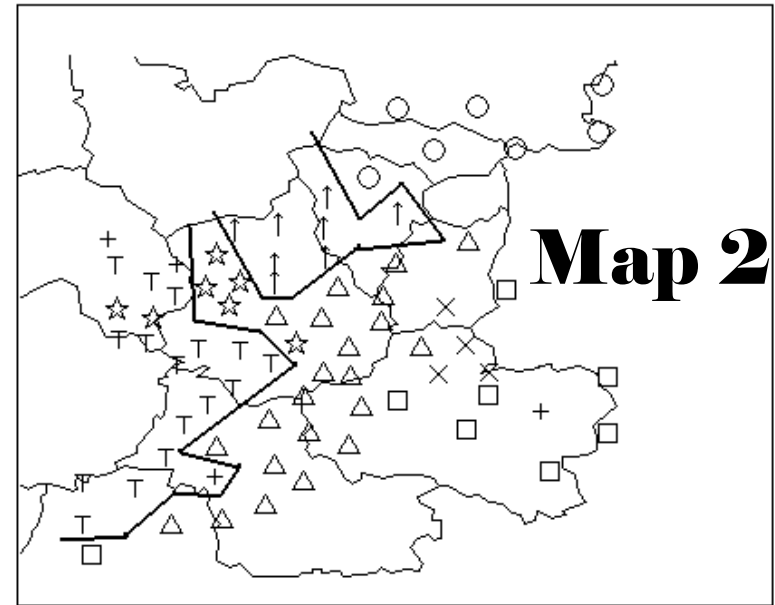
Standardized
points

Cluster ↑ + T X



Area of phonological variants

CASE 2
Complete Linkage Method
Manhattan Distance



Map 2

Ordinal scale 1-4 or 5

Cluster ○ □ △ ☆

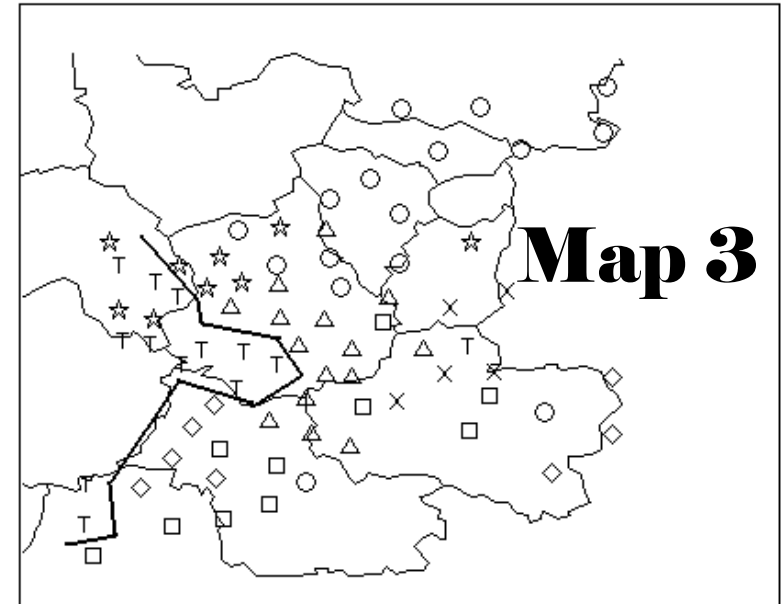
extended
standardized
zone

Cluster ↑ + T X

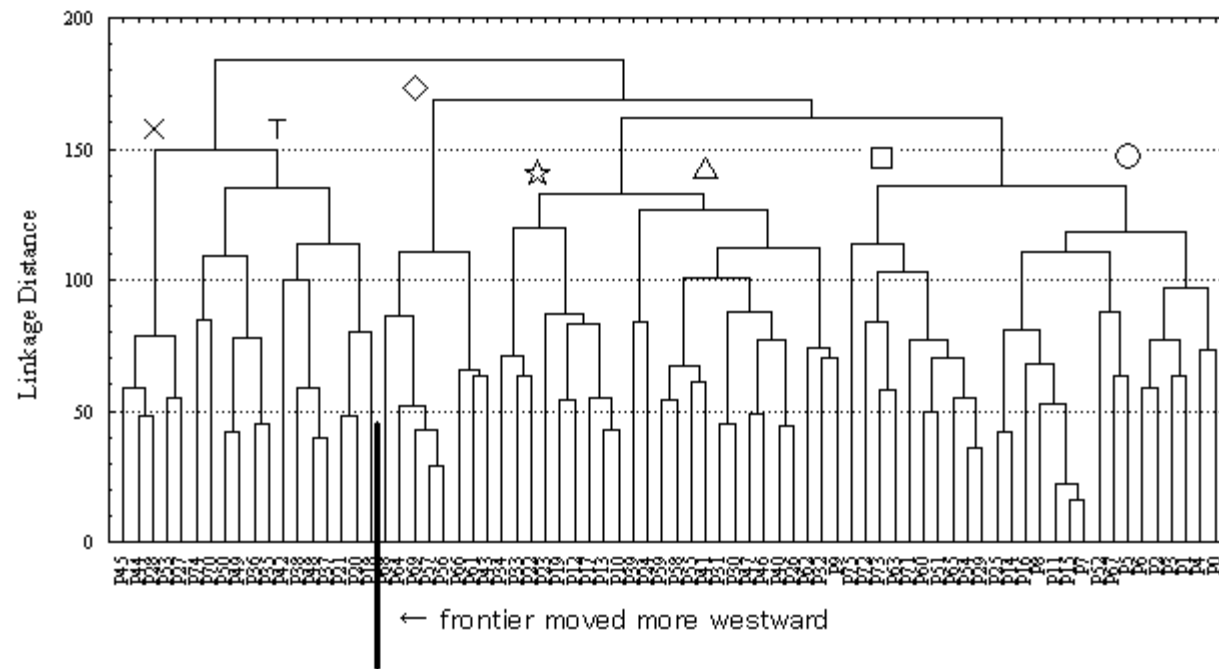
← frontier moved westward

Area of lexical variants

CASE 3
Complete Linkage Method
Manhattan Distance



Ordinal scale 1-10 with a special weighting



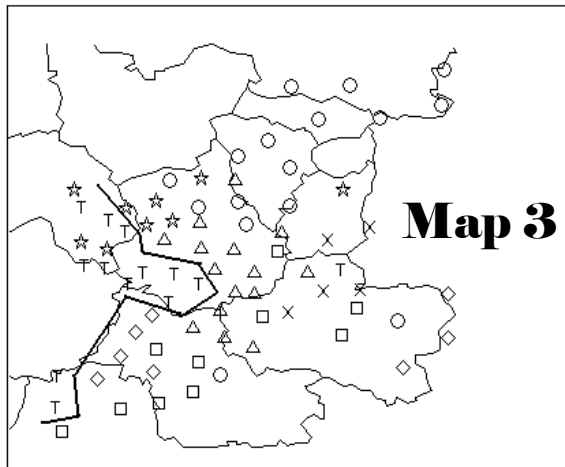
Cluster ○ □ △ ☆ ◇

Cluster T X
Isolated as area
of lexical variants

4. Can multivariate analysis contribute to the definition of the linguistic distance ?

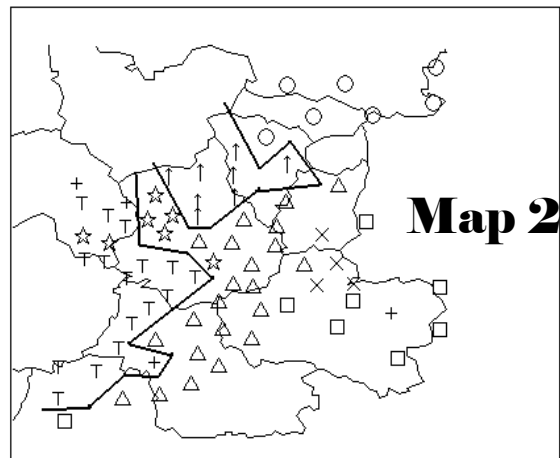
Conclusion of ALIFO data

CASE 3



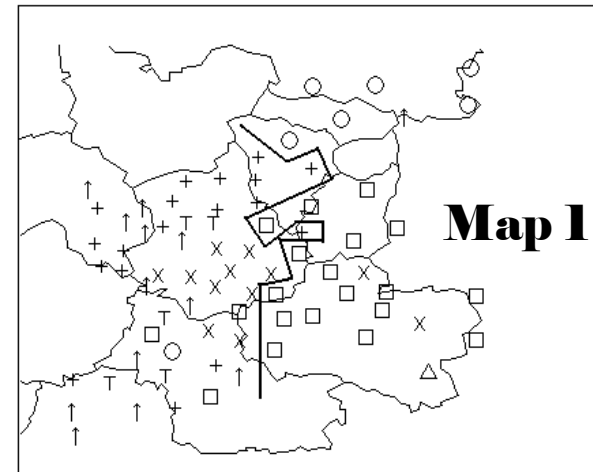
Ordinal scale 1-10 with a special weighting

CASE 2



Ordinal scale 1-4 or 5

CASE 1



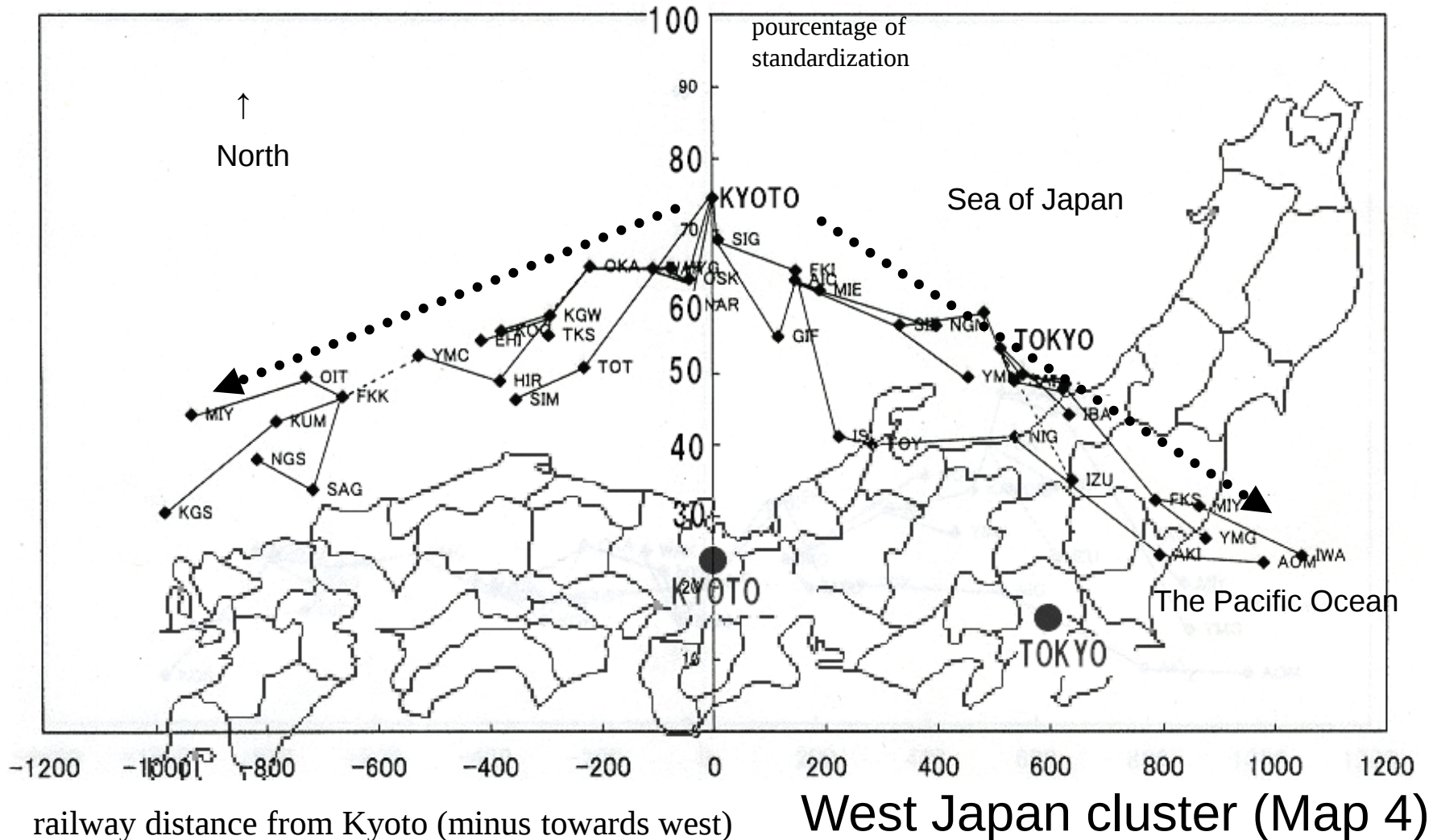
Nominal scale 1 or 2

area of lexical variation	area of morpho-phonological variants	area of standard form
T X	△ ☆	○
far ← → near		

5. Is it possible to use real distance ?

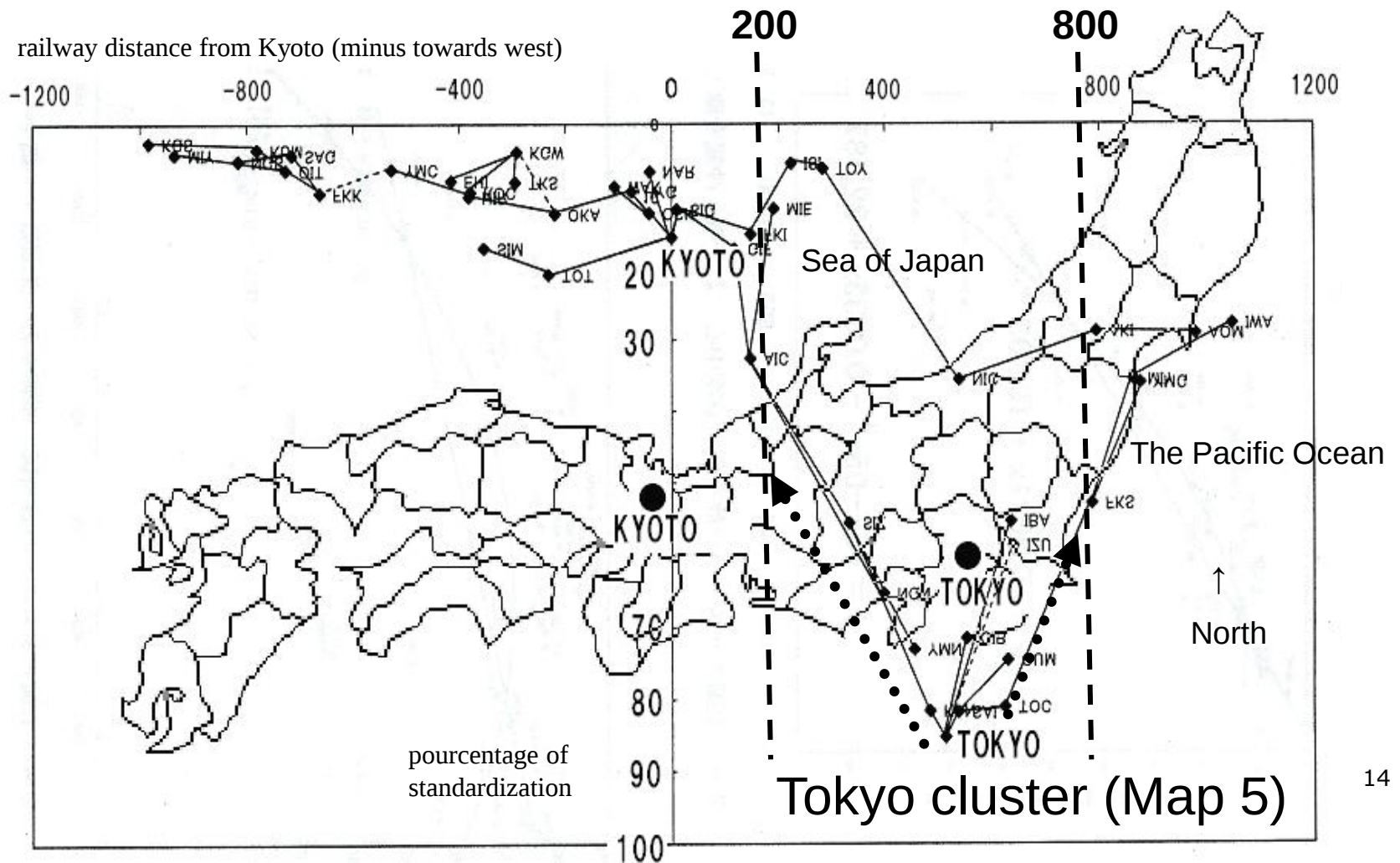
Standardization and railway distance

Hokkaido and Okinawa excepted, Inoue (2004) p.49 modified



Standardization and railway distance

Hokkaido and Okinawa excepted, Inoue (2004) p.53 modified



Conclusion

1. Efficiency of multivariate analysis even in the regions where geographical variants are hardly differentiated.
2. Significance of the distinction between morpho-phonological variants and lexical variants
3. Possibility of using the real distance like railway distance

References

- ◆ALIFO (*Atlas linguistique et ethnographique de l'Île-de-France et de l'Orléanais*), Marie-Rose Simoni-Aurembou : 1966, vol.1, 1969, vol.2, 1978, vol.3.
- ◆Eloy Jean Michel. 2004: "Des langues collatérales: problèmes et propositions", in: Eloy, J.-M. (ed.): *Des langues collatérales. Problèmes linguistiques, sociolinguistiques et glottopolitiques de la proximité linguistique*, vol. I, L'Harmattan, 5-25.
- ◆Everitt, B.S. 1993: *Cluster Analysis*, Third Edition, Edward Arnold, London/Melbourne/Auckland.
- ◆Goebel, Hans. 2002: "Analyse dialectométrique des structures de profondeur de l'ALF", *RliR*. 66: 5-63.
- ◆Goebel, Hans. 2004: "Bref aperçu sur les problèmes et méthodes de la dialectométrie (avec application à l'ALF)", in: Eloy, J.-M. (ed.): *Des langues collatérales. Problèmes linguistiques, sociolinguistiques et glottopolitiques de la proximité linguistique*, vol. I, L'Harmattan, 39-60.
- ◆Inoue, Fumio. 2004: "Three Stages of Diffusion of Standard Japanese Forms: Railway Distance and Percentage of Standard Usage by 4 Clusters" (in Japanese), *Gengo Kenkyu* 126, 69-92.
- ◆Inoue, Fumio and Hisako Kasai. 1989: "Dialect classification by standard Japanese forms", in: C. Mizutani (ed.), *Japanese Quantitative Linguistics, Quantitative Linguistics* 30, 220-235.
- ◆Kawaguchi, Yuji and Fumio Inoue. 2002: "Japanese Dialectology in Historical Perspectives", *Nouveaux regards sur la variation diatopique*, n° thématique conçu par M.R. Simoni-Aurembou, *Revue belge de philologie et d'histoire* 80, 801-829.
- ◆Martinet, André and Henriette Walter. 1973: *Dictionnaire de la prononciation française dans son usage réel*, France-Expansion.
- ◆Romesburg, H. Charles. 1989: *Cluster analysis for researchers*, Robert E. Krieger Publishing Company, Inc.
- ◆Yarimizu Kanetaka, Kawaguchi Yuji and Ichikawa Masafumi. 2005: "Multivariate Analysis in Dialectology – A Case Study of the Standardization in the Environs of Paris-", *Linguistic Informatics, -State of the Art and the Future- The first international conference on Linguistic Informatics*, Series Usage-Based Linguistic Informatics, Yuji Kawaguchi et al. eds., John Benjamins: 99-119.