

琉球大学学術リポジトリ

琉球列島の言語地図作成システム

メタデータ	言語: 出版者: 琉球大学工学部 公開日: 2007-08-23 キーワード (Ja): キーワード (En): Linguistic map, Ryukyu dialect, Distance between languages, Personal computer, Publication 作成者: 高良, 富夫, 北山, 敏昭, 宮城, 壮一, 屋慶名, 剛, Takara, Tomio, Kitayama, Toshiaki, Miyagi, Soichi, Yakena, Takeshi メールアドレス: 所属:
URL	http://hdl.handle.net/20.500.12000/1464

琉球列島の言語地図作成システム

高良 富夫* 北山 敏昭** 宮城 壮一+
屋慶名 剛**

A Linguistic Map Drawing System for the Ryukyu Dialect

Tomio TAKARA*, Toshiaki KITAYAMA**, Soichi MIYAGI+, and Takeshi YAKENA**

Abstract

In the Ryukyu dialect, we can easily find many words which have the same form of ancient Japanese words. It is said that geographical distribution of the word forms of the Ryukyu dialect represents the history of the Japanese language, i.e., the farther an island is from Shuri, the older form of the Japanese word is spoken in the island. Therefore, the linguistic map of the Ryukyu dialect gives an effective method to investigate the history of the Japanese language.

Okinawa Linguistic Research Center have inquired into and collected over 300 basic words spoken in over 800 of all villages of the Ryukyu islands. However, the collected data have not processed yet because the data involve of over 200,000 words.

In order to use these data effectively, we developed a system, on a personal computer, which can print the linguistic map of the Ryukyu dialect. When we select a certain word, the system prints the word forms of each village at each point of the village on the map. This linguistic map is shown on a computer display and can be printed out with high quality available for the publication. The system can also show similarity or difference of the dialects by calculating the distance between the dialects using distances between words and between phonemes.

Key Words: Linguistic map, Ryukyu dialect, Distance between languages, Personal computer, Publication

1. まえがき

琉球方言は、琉球列島の各地で話される方言の総称であり、本土方言とともに日本語を二分する大方言である。方言の語形は地域により大きく異なっており、琉球方言の中には古代日本語の特徴を色濃く残しているものも多い。琉球方言では、首里から離れた地域の方言ほど古い語形を保っており、琉球方言の地理的分布は日本語の歴史を表すとも言われる。このようなことから、琉球方言の言語地図は、日本語の歴史を研するうえでも有効な手段を与える。

沖縄言語研究センターは、1979年4月から12年にわたって、「琉球列島の言語の研究」[1],[2]として、方言の臨地調査を行ってきた。この調査は、奄美から与那国まで、南北1000kmに及ぶ琉球列島に分布する多様な方言を研究し、古い日本語の特徴を含むといわれる琉球方言の生成・発展の特徴を調べるために行われた。その方言資料は、琉球列島全体で約800に及ぶすべての古い村落を対象にして、統一された調査票を用いて300あまりの基本語を中心に調査されたものである。しかし、これらの資料は20万件以上の膨大なものであるため、現在まで、その整理が必ずしも十分なされているとはいえない。

そこで本研究では、この方言資料を有効に利用するため、これを方言地図として出力できるシステムを開発した。本システムでは、ある単語の各地域における語形を地図上の各地点に表示する。これは、コンピュータの画面に表示されるとともに、出版にも耐える程度の高品質な方言地図として印刷することができる。また方言間の相違度を方言間距離として数値化して示すことができる。第2章では、言語地図作成システムの概要を述べ、このシステムで作成された方言地図の例を示す。第3章では方言間距離について述べ、首里を中心

受理： 1995年11月6日
* 工学部情報工学科
(Dept. of Information Eng., Fac. of Eng.)
** セコム琉球(株)
(Secom Ryukyu Co.)
+ 国際システム(株)
(International Systems Co.)
++ 工学部電子・情報工学科
(Dept. of Electronics & Information Eng., Fac. of Eng.)

沖縄言語研究センター研究発表会において1994年6月5日および1995年7月22日に発表済み

に計測した方言間距離の例を示す。第4章は、本システムのマニュアルであり、第5章では、さらに検討すべき課題を述べる。

2. 言語地図システムの概要

本システムは、以下のハードウェアにより構成されている。
 コンピュータ本体: Apple Macintosh Quadra 840AV
 ディスプレイ: Apple Macintosh 21" Color Display
 デジタイザ: MUTOH IS/ADB
 レーザ・プリンタ: OKI MICROLINE 801PSII

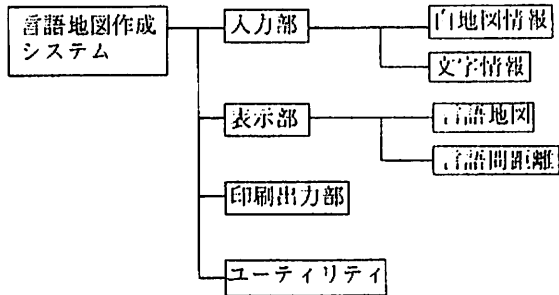


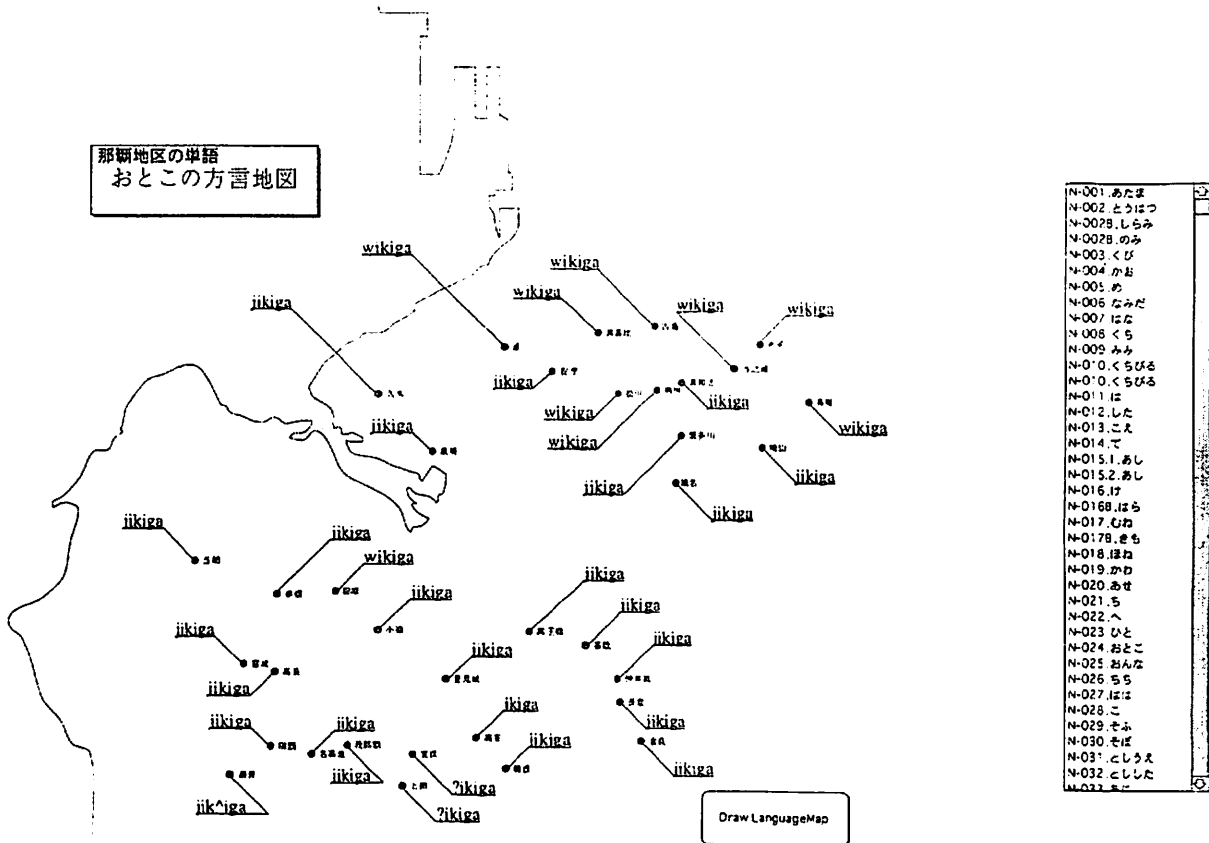
図1. 言語地図作成システム

本システムを構成するために使用したソフトウェアは以下のとおりである。

システム本体の作成: Claris HyperCard J-2.2.1
 Zedcor FutureBasic
 白地図作成: Aldus FreeHand 3.1J
 Adobe Photoshop 2.5J
 ユーティリティ: YooEdit 0.96b3
 音楽記号作成: ALTSYS FONTOGRAPHER

システム構成を図1に示し、以下にその概要を説明する。

- (1) 入力部: システムに白地図や方言データなどを入力する。これにより、新たな地域の方言地図を作成し、システムを拡張することができる。白地図はデジタイザを用いて作成する。方言データなどの文字情報は、沖縄言語研究センターからフロッピーディスクなどにより提供される。
- (2) 表示部: 表示部には言語地図と言語間距離の2つがある。言語地図の表示はHyperCardで行われる。言語間距離部はFutureBasicで作成されたプログラムで実行される。
- (3) 印刷出力部: 言語地図の印刷出力を行う。HyperCardでは高品質の地図を印刷出力することができないので、本システムでは、白地図をFreeHandで印刷出力し、その白地図の上にHyperCardで文字情報を印刷する。
- (4) ユーティリティ: システムに入力される方言データは、



通過ルート → 那覇島 → 沖縄本島 → 那覇の方言地図

図2. 方言地図の出力例 “おとこ” (那覇市周辺)

NEC社のパソコンPC9801RA上でデータベース・ソフト F1DATABOXを使って作成しているため、そのままではデータの構造上、本システムでは扱いにくい。そこで、データの構造を変更するためのユーティリティを用意した。また、音素記号を新たに作成・登録することにより、今までにない音素記号も扱うことができる。

本システムの操作の概要は以下の通りである。まずシステムを起動するとタイトル画面が現れる。そこでマウスのボタンを押すと琉球列島の地図が現れる。ここで方言地図を表示したい島または地域をマウス・カーソルで選択すると、その地域が拡大される。方言地図表示ボタンをクリックし、単語を選択すると、画面の各地点にその単語の方言の形が表示される。例として、沖縄本島を選択し、さらに那覇市周辺を選択して表示した方言地図を図2に示す。

3. 方言間距離

言語間距離とは、言語間の相違度を数値で表したものである。例えば、2つの集落の方言間で、発音の異なる単語の数または異なる単語の割合が方言間距離である。言語間距離をさらに厳密に求めるためには、単語の発音の相違度を規定する必要がある。単語間の相違度を単語間距離と呼ぶ。例えば、単語に含まれる音素が完全に一致した場合を同じとみなすのか、いくつかの音素が一致すれば同じとみなすのか、などである。さらに、音素の一致性については、類似の発音の音素とそうでないものを区別し、類似性を数値化する必要がある。これを音素間距離と呼ぶ。

言語間距離を定義するため、単語間距離および音素間距離を導入し、これらを研究者の観点により変更できるようにすれば、言語間距離を計測した結果を考察する上で有用である。

以下に本システムで使用している音素間距離、単語間距離、言語間距離について述べる。

(1) 音素間の距離

音素とは単語の構成要素のことである。本システムではChomsky&Halleの弁別的素性[3]を利用して音素間の距離を求める。弁別的素性を表1に示す。表1において+は、その性質を持っていることを表し、-は逆の性質を持っているこ

とを表している。0はChomsky&Halleの表で空白の部分を表す。なお表1において?, ʀ, N, ɹ, j, ɿはChomsky &Halleの表にないので、著者らが暫定的に弁別的素性を設定した。

本システムでは弁別的素性の各要素に次のような値を対応させ、音素を弁別的素性を成分とするベクトルとみなす。

- 記号が+のとき: 1
- 記号が-のとき: -1
- 記号が0のとき: 0

本システムではこの弁別的素性を用いて音素間の距離を定義する。ある2つの音素について音素ベクトルの各成分の差を求め、各成分の絶対値の和を音素間の距離とする。例えば音素 a と音素 i の場合は次のようになる。

音素 a は(1,-1,-1,1,1,-1,-1,-1,0,0,0,0), 音素 i は(1,-1,1,-1,-1,-1,-1,-1,0,0,0,0)のような素性を持っている。したがって音素間距離は次式より8となる。

$$|0|+|+0|+|2|+|2|+|2|+|0|+|0|+|0|+|2|+|0|+|0|+|0|+|0|=8$$

この値が小さいほど比較した2つの音素が似ているといえる。ここで音素間距離にしきい値をもうける。音素間距離がしきい値以下なら比較している2つの音素は一致しているとみなし、しきい値よりも大きい場合は一致しないとみなす。この音素の一致・不一致を規定するしきい値は、研究者の観点により、任意に変えることができる。しきい値を大きくすると、音素対応法則において対応する音素は同じ音素とみなすこと[4]ができる。

このようにして得られた音素間距離は、単語間の距離を求めるDPマッチングで利用される。

(2) 単語間の距離

本システムでは単語間の距離を求めるためDPマッチング法[5]を用いる。DPマッチングでは比較する単語の一致度を調べることができる。

DPマッチング法を簡単に説明する。s₁ s₂ … s₁ … s_mをある文字列(単語)とし、t₁ t₂ … t₁ … t_nを他の文字列とする。この2つが最もよくマッチしたときのずれの量 f(i, j)を導入し、これを次のような漸化式によって計算する。

表1 弁別的素性

	a	ĩ	i	ı	u	e	ɜ	o	j	w	r	p	b	m	ŋ	t	d	N	n	ɲ	s	ʃ	z	ʒ	c	k	g	h	ϕ	q	?					
母音性	+	+	+	+	+	+	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-					
子音性	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0	+				
高性	-	+	+	+	+	-	-	-	+	+	-	-	-	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	0	-				
後性	+	-	-	-	+	-	-	+	-	+	-	-	-	-	-	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	0	+				
低性	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	0	-			
前方性	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-			
舌頂性	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	0	-		
円唇性	-	-	-	-	+	-	-	+	-	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	
緊張性	+	-	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
有声性	0	0	0	-	0	0	+	0	0	0	+	-	+	+	+	-	+	+	+	+	-	-	+	+	-	-	+	-	-	-	-	0	-			
持続音性	0	0	0	0	0	0	0	0	0	0	+	-	-	-	-	-	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-	+	+	0	-	
鼻音性	0	+	0	0	0	0	-	0	0	0	-	-	-	+	+	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-	
粗擦音性	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	0	-

$$\begin{aligned} \text{初期条件} \quad & f(0, 0) = 0 \\ & f(i, 0) = \infty \quad (i = 1 \sim m) \\ & f(0, j) = \infty \quad (j = 1 \sim n) \end{aligned}$$

$$f(i, j) = \min\{f(i-1, j)+1, f(i, j-1)+1, f(i-1, j-1)+d(s_i, t_j)\}$$

$$d(s_i, t_j) = \begin{cases} 0 & (\text{dis}(s_i, t_j) > \theta) \\ 1 & (\text{dis}(s_i, t_j) \leq \theta) \end{cases}$$

ここで、関数 dis は音素 s_i と t_j との音素間距離である。 θ は音素一致の基準を定める値であり、「音素間距離」で述べたように、この値を可変値にして一致の基準をゆるめることができる。

このようにしてマッチングを行なった結果は、比較した2つの単語で不一致である音素の数(挿入と削除を含む)を示している。

ここでも音素間距離のときと同様に、しきい値をもうけ、ある単語における不一致の音素の数がしきい値以下なら、その単語は一致とみなし、しきい値より大きければ不一致とみなす。しきい値は、研究者の観点により変更できる。他の研究[6]では、単語の最初の子音が一致しただけで単語が一致したとみなしている場合がある。

(3) 言語間の距離

言語間の距離は次のように定義する。

- (i) 2つの集落間で不一致とみなされた単語の数を求める。
- (ii) 2つの集落間で比較した単語の総数を求める。
- (iii) 上記の(i)を(ii)で割る。

すなわち言語間距離は、2つの集落における比較単語の総数に対する不一致の単語の数の割合である。

(4) 計測例

本システムを用いて言語間距離を求めた結果の例を図3に示す。ここでは那覇市首里を中心にした言語間距離を求めた。

首里を中心にした言語間距離では、金武町と読谷村の間を境にして沖縄本島の方言が大きく2つに分類されることが分かる。沖縄本島の方言は大きく沖縄北部方言と沖縄中南部方言とに分けることができ、その境界は恩納村、金武町の南側の市町村境界線上にあるということは、方言学では、すでに定説である。言語間距離の計測結果は、この方言学の定説を数値的に実証したものといえる。

4. 言語地図作成システムのマニュアル

本システムの構成を図1に示す。本システムは、入力部、表示部、印刷出力部、ユーティリティから成る。以下に各部について詳細に述べる。

4.1 入力部

【言語地図用方言データ】

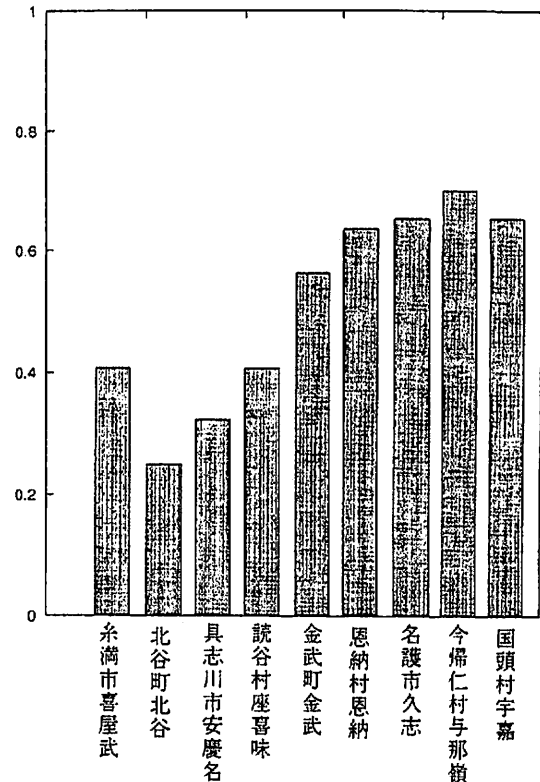


図3. 言語間距離の測定結果 (那覇市首里中心)

音素間距離しきい値: 4

単語間距離しきい値: 1

沖縄言語研究センターから受け取った「方言のデータ」を、本システムの方言地図用データとしてシステムに組み込むまでの手順を、図4の流れにしたがって説明する。

- (1) 沖縄言語研究センターからは以下のような条件で方言データが提供される。

メディア: 3.5インチ, 2HD, 1.2MBのMS-DOS
フォーマット

ファイル形式: TEXT

ファイル名: 集落名

データの区切り: コンマ

- (2) Macintoshでは直接1.2MBのMS-DOSフォーマットのフロッピーは読めないので、まずPC-9801で読み込んで、これをPC-9801のftpコマンドを使ってUNIXワークステーションsw021に転送する。

- (3) sw021に送られたデータは自動的に文字コードがS-JISからEUCに変換されるので、sw021上でnkfコマンドを使ってS-JISに戻す。nkfコマンドの書式を以下に示す。

nkf -s 変換前のファイル名 > 変換後のファイル名

- (4) sw021でコード変換したデータを、Macintoshの通信ソフト「NCSA telnet 2.5J2」のftpコマンドを用いてMacintoshの「方言地図用データ」のフォルダに転送する。「方言地図用データ」のフォルダは以下に示すディレクトリに存在する。

Macintosh HD : 方言地図 : 方言地図用データ

<沖縄言語研究センター>

<高良研究室>

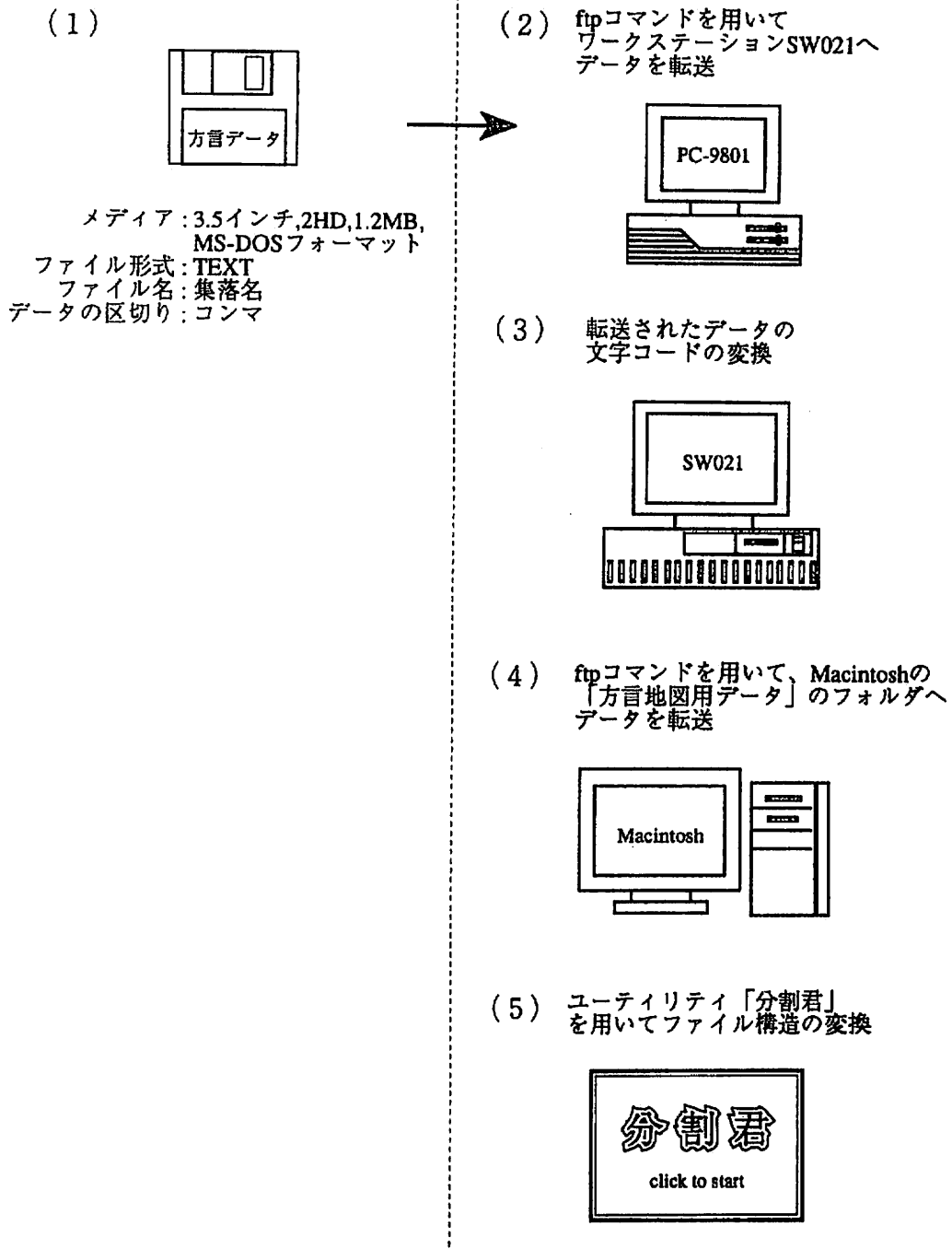


図4. 方言のデータをシステムに組み込むまでの流れ

- (5) Macintosh上でユーティリティ「分割君」を使ってファイルの構造を変換する。「分割君」の使い方は、4.4節のユーティリティで示す。

【白地図データ】

白地図データの作成とシステムへの組み込み方を順を追って説明する。言語地図用の白地図データとしては、「画面表示用」と「印刷出力用」の2つが必要である。

- (1) デジタイザ (MUTOH社のIS/ADB) をMacintoshに接続する。
- (2) デジタイザの接続設定を行うため、Macintoshを再起動する。
- (3) Aldus FreeHand 3.1Jを起動する。
- (4) デジタイザの入力器具 (クロスカーソル) を用いて目的の地図をトレースする。このとき、後で画面表示用白地図に集落名を入力しやすいように各集落の場所に目印を付けておく。
- (5) 地図を印刷する。
- (6) ファイルメニューの「データ書き出し」で「Mac用EPS」を選択し、適当なファイル名を付けて保存する。このとき、保存するファイル名の最後に「表示用」と書き加える。ここで保存したファイルは、「画面表示用」に用いる。
- (7) ファイルメニューの「保存」を選択し、適当な名前をつけて保存する。このとき、保存するファイル名の最後に(印刷用)と書き加える。ここで保存したファイルは(19)で地図の大きさを調整をした後、「印刷出力用」として用いる。
- (8) FreeHandを終了する。
- (9) Adobe Photoshop 2.5Jを起動する。
- (10) ファイルメニューの「指定形式で開く」を選択すると、ダイアログボックスが出てくるので、(6)で「表示用」として保存したファイルを開く。このときダイアログボックスの「ファイル形式:」のところは「EPS」にする。
- (11) さらにダイアログボックスが出てくるが、無条件に「OK」をクリックする。
- (12) Photoshopの画面に表示されている地図を選択範囲メニューの「全画面の選択」で選択する。
- (13) 編集メニューの「コピー」で地図をコピーする。ここでコピーした地図は(18)においてペーストして、画面表示用の地図として使う。なお、この画面表示用の白地図のファイルは、バックアップ用として残しておく。
- (14) Photoshopを終了する。
- (15) 方言地図表示システムを起動する。HyperCardJ-2.2.1は、これと同時に自動的に起動される。
- (16) 編集メニューの「新規カード」を選択し、新たにカードを作る。このカード上に新たな地域の方言地図を作成していく。
- (17) オブジェクトメニューの「新規バックグラウンド」を選択し、(16)で作成したカードに新たなバックグラウンドを作成する。(注: 新規バックグラウンドを作成すると、自動的にバックグラウンド編集状態になる。)
- (18) この状態で、編集メニューの「ペースト」を選択し、

(13)でコピーしておいた地図をバックグラウンドにペーストする。

- (19) 編集メニューから「バックグラウンド」を選択し、フォアグラウンドに戻る。
 - (20) ファイルメニューから「カードプリント」を選択し、カードを印刷してみる。
 - (21) (5)で印刷した地図と(20)で印刷した地図の大きさを比べ、(5)の地図の大きさを(20)で印刷した地図と同じ大きさになるように(5)の地図をFreeHandで調整して保存する。これで、印刷出力用の白地図が完成する。
 - (22) (18)でペーストされた白地図上の、各集落の位置を示した目印に重ねて直径約2[mm]くらいの黒丸を描く。
 - (23) 黒丸のすぐ右に集落名を書き込む。
 - (24) 黒丸から右上45度の方向に、適当な長さの引出線を描く。
- 以上で、方言地図表示部へ白地図を入力する操作は完了する。

【方言間距離用データ】

方言間距離を計算するためのデータは、「言語地図用方言データの入力」の(1)~(4)までの手順とほぼ同様にしてMacintoshに読み込む。異なる部分を次に示す。

- (1) 方言間距離用データの格納先は「方言間距離用データ」フォルダとする。そのディレクトリを以下に示す。
Macintosh HD : 方言間距離 : 方言間距離用データ
- (2) 「分割君」を用いたファイルの構造変換を行わない。方言地図表示部と方言間距離部で用いる方言のデータの内容は全く同じものである。しかし、方言地図表示部では、検索の高速化のため特別なファイル構造に変換する。距離計算のためには、変換を行う前のデータを別に持たなければならない。

4.2 表示部

【方言地図】

方言地図の表示部はHyperCard J-2.2.1で作成されており、「言語地図作成システム」の本体である。ここでは、白地図上にそれぞれの集落における方言形を記した「方言地図」を描く。以下に、那覇市の方言地図を描くまでのシステムの操作方法を順を追って説明する。

- (1) 「言語地図作成システム」のアイコンをダブルクリックすると、システムが起動する。
- (2) 「click to start」をクリックすると、琉球列島の地図が表示される。
- (3) 沖縄本島をクリックすると、琉球列島の中の沖縄本島が拡大して表示される。
- (4) 地図中の「那覇地区」をクリックすると、那覇地区が拡大して表示される。
- (5) 地図中の「Draw LanguageMap」のボタンをクリックする。
- (6) 「描きたい方言地図の単語を選んでください」というメッセージとともに「OK」ボタンが出てくるので、それをク

リックする。

- (7) 画面の右端に共通語の単語一覧表が表示されるので、その中の単語をマウスでクリックして選択すると、その選択された単語の方言地図が描かれる。
- (8) 別の単語の方言地図を描きたいときは、(7)と同様に単語一覧表からその単語を選択しクリックする。単語一覧表には一度に39単語しか表示できないが、右端についているスクロールバーを使って表をスクロールすることにより、371単語全てを見ることができる。
- (9) システムを終了するには、ファイルメニューから「HyperCard終了」を選択する。

図 2 に、言語地図作成システムの画面出力の例が示されている。

【方言間距離】

ここでは集落間の方言の差異を具体的な数値として求めることが出来る。以下に、方言間距離を求めるまでの操作法を示す。

- (1) 「方言間距離」のアイコンをダブルクリックして起動する。
- (2) ダイアログボックスが出てくるので、各集落の方言のデータが置かれたフォルダの中から任意の集落Aのファイルを開く。(注：ファイル名は集落名と同じになっている。)
- (3) 再びダイアログボックスが開くので、(2)で選択した集落Aとの間で方言間距離を求めたい集落Bのファイルを開く。
- (4) 「hougenkankyori=」の後ろに集落Aと集落Bの方言間距離が表示される。
- (5) 【方言間距離】を終了するには、ファイルメニューから「Quit」を選択する。

4.3 印刷出力部

方言地図の印刷出力をする。HyperCardでは、目の粗いビットマップの印刷出力だけができる。そのため、地図の輪郭線がギザギザになり綺麗な印刷結果が得られない。

そこで、見栄えのする方言地図を印刷するため、少し手間はかかるがAldus社のFreeHandで白地図部分を印刷しておき、その上にHyperCardで音素記号や集落名を重ねて印刷するという方式を取っている。以下に、印刷出力の手順をより詳しく示す。

- (1) 4.1で説明した「印刷出力用の白地図」ファイルをFreeHandで印刷出力する。
- (2) 「言語地図作成システム」を起動する。
- (3) 印刷出力したい画面を表示させる。
- (4) 「言語地図作成システム」では、方言地図の白地図をカードのバックグラウンドに描かせているので、編集メニューで「バックグラウンド」を選択し、バックグラウンドに移動する。
- (5) ツールメニューで「選択ツール」を選択する。
- (6) ペイントメニューで「すべてを選択」を選択する。
- (7) 編集メニューで「カット」を選択する。このとき「消去」ではなく「カット」を必ず選択する。

- (8) 再び、編集メニューから「バックグラウンド」を選択することによりフォアグラウンドに戻る。このとき白地図が消えて、音素記号と集落名およびそれらを結び引出線だけが表示されているのを確認する。
- (9) (1)で印刷しておいた白地図を再びプリンターにかけ、(8)の白地図の消された方言地図を印刷出力する。
- (10) 印刷が終了したら、再び「バックグラウンド」に戻り、編集メニューから「ペースト」を選択する。すると(7)で「カット」した白地図が再び表示される。(注：もしも、(7)で白地図を消すときに「カット」ではなく「消去」で消すと、ここで「ペースト」をしても白地図が表示されない。すなわち、白地図が消滅するので注意が必要である。もし、あやまって消去してしまった場合は、4.1の【白地図データ】の(6)で保存したバックアップファイルからコピー&ペーストをする。)

4.4 ユーティリティ

【音素記号フォントの作成】

本システムでは、沖縄言語研究センターで入力された方言のデータを利用している。しかし、本システムの開発環境がApple社のMacintoshであるのに対して、沖縄言語研究センターでは富士通のワープロでデータ入力を行っている。そのため琉球方言を表現するための特殊な音素記号のフォントを共有することが出来ず、直接データの受け渡しをすることが出来ない。そこで、特殊な音素記号を普通のアスキー文字に置き換えて、データの受け渡しを行うという方法をとっている。どの音素記号をどの文字に置き換えているかを示した対応表を表2に示す。

Macintoshで方言データを見るときは、フォントを「HatsuoKigouFont」にする。このフォントは、本研究室においてALTSYS社のFontoGrapherを用いて著者らが作成したものである。

注意しなければならない例外について説明する。

Macintoshで特殊なフォントを作成するために用いているFontoGrapherは、扱える文字数に限りがある。そのため、(コントロールキー) + (文字キー) で表す文字までも使わなければならない。しかし、この(コントロールキー) + (文字キー) は、沖縄言語研究センターのワープロでは入力することができない。そこで沖縄言語研究センターでは、代わりに全角のアルファベットで入力してもらい、Macintosh上で、(コントロールキー) + (文字キー) に置換するという操作をすることになった。今後、扱わなければならない特殊な音素記号が増えた場合、すべてこの方法で対応するほかない。

【方言地図用データの変換】

言語地図表示部はHyperCardで作成されている。しかし、HyperCardは一つ一つの操作が非常に遅いという欠点がある。そこで少しでもデータ検索のスピードを上げるため、一つの大きなファイルを多くの小さなファイルの集合に分割するという方法をとる。ファイルを分割するためのツールとして「分割君」を作成した。

「分割君」の使い方を以下に説明する。

表2 フォントの対応表

1	0	^	A	I	P	t				
#	1	v	B	l	Q	Q		A	^a	i
%	2	r	C	ç	R	э		B	^b	i
&	3	`	D	d	S	ş		C	^d	i
'	4	l	E	ε	T	η		D	^e	N
@	5	†	F	f	U	U		E	^f	t
©	6	-	G	g	V	u		F	^g	p
?	7	-	H	h	W	w		G	^j	m
!	8	-	J	l	X	θ		H	^l	d
q	9	-	I	p	Y	i		I	^n	'
~	:		K	z	Z	3		J	^o	'
™	.		L	i	J	i		K	^u	k
†	<	ü	M	u	J	i		L	^v	s
!	=	w	N	N				M	^w	i
	>	o	O	φ						

- (1) 初めに、全ての方言データのファイル名を、集落名から数字の番号に変更する。この操作は、ファイル名と白地図上の各集落の方言データを表示するフィールドの番号を同じにするために行う。
- (2) ファイル名(数字)と同じ名前のフォルダーを作成し、その中に、さきほど名前を変更した方言データのファイルを入れる。
- (3) 「分割君」を起動する。
- (4) スタートボタンをクリックする。
- (5) ダイアログボックスがでてくるので、分割したいファイルを選択する。

以上の操作で言語地図表示部で扱うことができるデータ構造に変換することが出来る。なお、このように分割したデータは、方言間距離を求めるためには使えない。

5. 検討

現時点までに作製された言語地図作成システムは、必ずしも十分なものではない。今後検討しなければならない課題を以下に述べる。

まず、問題点として、本システムを使用したとき、一つ一つの動作が非常に遅いということがあげられる。例えば、単語の一覧表からひとつの単語を選択し、その単語の言語地図が描かれるまでには15秒を要する。これは、データ構造に問題があると考え、データを分割するなどの工夫をしたが改善されていない。今後さらにデータ構造などを検討する必要がある。

方言学の専門家による本システムの使用試験を行った。ここで指摘された改良すべき点を以下に示す。

- (1) 異なる方言形ごとに色分けして示す。多数の地点における方言形は文字だけでは区別しにくい。これを同じ方言形は同じ色にするなど色分けを行うと直感的に分かりやすい。
- (2) 方言形の異なる文字だけを示す。方言形は、同じ意味の単語はほぼ同じ音素列で表され、異なるのは1, 2音素のみ

である。異なる文字だけを表示すれば違いの分布が直感的に分かりやすい。

- (3) 白地図では、市町村の境界も破線などで表示する。“間切り”の境界も表示できれば、歴史的考察も容易になる。
- (4) 沖縄本島を数カ所に分割した地域の方言地図や沖縄本島全体の方言地図など、より広範囲の方言地図とする。
- (5) 広範囲の方言地図では表示すべき地点数が多くなるので、方言形ごとに○△などの記号で表示したほうが分かりやすい。

本システムでは、音素間距離を定義するため、音素の弁別的素性を用いたが、琉球方言で使用される音素の中には、Chomsky & Halle の表にはない音素もある。本システムでは、これらの音素には、類似の音素から筆者らが推定した弁別的素性を与えている。さらに正確な方言間距離を得るためには、琉球方言特有の音素に、より正確な弁別的素性を与える必要がある。この方法として以下のことが考えられる。

- (1) 音声学の権威にこれを依頼する。
- (2) 国際音声記号の各音素の特徴[7]と弁別的素性との対応から推定する。
- (3) (2)をニューラルネットにより実行する。

本システムでは単語間距離を計算するとき、音素間距離を基にDPマッチングを実行している。しかし、日本語では、音素よりモーラが基本単位であると考えられる。モーラを単位としたDPマッチングに改良する必要がある。

方言間距離を計算するとき、全ての単語間距離を同等にあつかっている。しかし、単語の使用頻度は単語ごとに異なるから、これは方言間の相違度としては十分でない。単語の使用頻度に応じて重みを付けた距離を用いる必要がある。また、単語を人体に関する用語、動物に関する用語、植物に関する用語などに分類し、それぞれ別々に距離を測定することも有用である。

6. むすび

沖縄言語研究センターが収集した膨大な方言資料を有効に利用するため、方言地図を作成するシステムをパーソナルコンピュータ上に構成した。本システムでは、特殊な音素記号を含む方言地図を高品質に印刷出力することができる。また、方言間の類似度を方言間距離として数量化して示すことができる。

今後の課題は、まず方言間距離を利用して、似ている方言は近くに、似ていない方言は遠くにあるように変形した地図を作成できるようにすることである。これにより、言語的な遠近関係と地理的な遠近関係との違いから、集落間の文化的な交流と遠近関係を容易に考察することができる。さらに、音素間距離、単語間距離、方言間距離のそれぞれを任意に定義・設定できるようにすることがあげられる。研究者の直感に合う距離がシステムに設定できれば、これはその研究者の思想を表現したものとなり、言語学の研究支援システムとして大いに役立つと考える。

文 献

- [1] 沖縄言語研究センター：“沖縄言語研究センター会報”，第1～14号，沖縄言語研究センター(1980-01～1992-06)。
- [2] 沖縄言語研究センター：“奄美諸島方言の言語地理学的研究”，沖縄言語研究センター資料，No.88，p.5 (1990-03)。
- [3] 長尾 真：“言語工学”，昭晃堂，p.3(1983)。
- [4] 高良富夫，久場長司：“日本語共通語と琉球方言との間の音素対応分析エキスパートシステム”，琉球大学工学部紀要，第39号，pp.99-108 (1990-3)。
- [5] 長尾 真：“言語工学”，昭晃堂，pp.72-74(1983)。
- [6] 安本美典，野崎昭弘：“言語の数理”，筑摩書房，pp.165-198(1976)。
- [7] 日本音声学会：“音声学大辞典”，三修社，pp.955-963 (1976)。