

PC-9801上の言語地図作成支援システム: EGL

前川 喜久雄
(鳥取大学)

要旨 EGLは16ビットパソコンPC-9801で稼働する言語地図作成支援システムである。編集機能に優れた大型機上のGLAPSと地図の表現力に優れたパソコン上のSEALの特長を両立させることを目的として開発された。EGLでは完全な対話型の操作によって、語形のまま入力されたデータの編集と言語地図の作成が可能である。地図の出力媒体は現在のところCRTのみであるが、将来的にはプロッタ等を予定している。語形編集終了後、地図作成に要する時間は364地点のデータを使用した場合で20秒を切っており、従来の手作業に対し平均して少なくとも100倍以上のスピードを実現している。

Editor for Geo-Linguistics : A Conversational System on NEC PC-9801
for the Analysis and Cartography of Geo-linguistic data

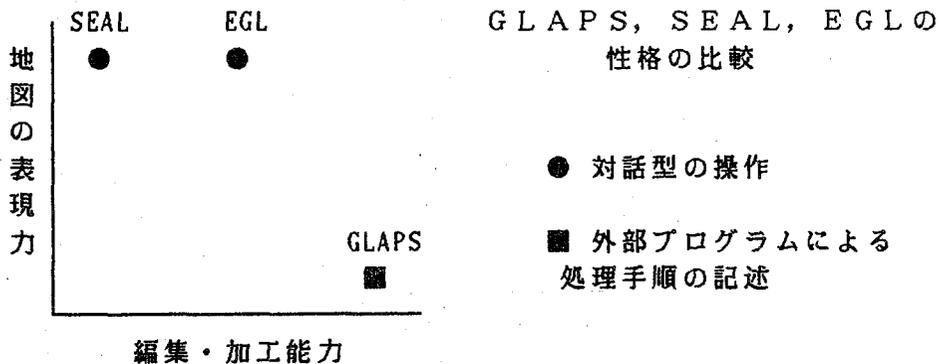
MAEKAWA Kikuo
(Tottori University)

Abstract : EGL is a conversational software system on NEC PC-9801 developed in order to minimize that enormous time and efforts required for the analysis and making up of linguistic maps. Based on the critical analysis of two predecessor programs GLAPS and SEAL, EGL is designed to incorporate both the editing power of GLAPS and the graphics output of SEAL. Though the output device of linguistic map now available is limited to CRT, plotter output will be served in the near future. The time needed to make a map on CRT is less than 20 seconds when Linguistic Atlas around Nakaumi and Daisen's 364 informants' data are used. It is a hundred times faster than the manual cartography at the very least.

1 EGLの目的

言語地図の作成は多くの労力を要する作業である。しかもその作業の多くの部分はハンコ（ゴム印）押しに代表される単純作業である。この作業から言語地理学に携わる者を解放し、地図の解釈などの本質的な作業に専念しやすい環境を提供するために単純作業をコンピュータに肩代りさせようとするのはごく自然な欲求である。今回発表するEGLと呼ぶプログラムもこの欲求を満たそうとするものである。

言語地図の作成を支援するコンピュータプログラムとしては荻野綱男氏のGLAPS、および福嶋秩子氏のSEALが既に知られている。この2種のプログラムの性格を(1)言語資料の編集・加工能力、(2)作成される言語地図の表現力、(3)操作の方法という三つの観点からごく概念的に紹介すると下図のようになる。表現力のSEALに対し、編集能力のGLAPSとすることができる。そしてこの性格は両者の開発環境を端的に反映している。GLAPSは大型コンピュータ上でのバッチ処理が普通であった時代に開発されたプログラムであり、SEALは初期の8ビットパソコンのために開発されている。GLAPSが地図の出力媒体としてラインプリンターを採用したことや、SEALがプログラムとデータを分離していないこと（言語データはBASICのDATA文として与える）等は共に時代の制約であったとすることができよう。



EGLの狙いは上のこの図にも明らかなように、GLAPSに近い編集能力を備え、SEALと同等の表現力を持ち、対話型の操作が可能なプログラムである。このうち操作方法として対話型とプログラム型のどちらが良いかは簡単に決められない問題である。ワープロのような対話型は初心者にとっつきやすい反面、決まりきった処理しかできないことになりかねないからである。EGLは全くコンピュータに触れたことのない学生にも使用させるつもりで開発したので対話型の操作を重視している。

2 EGLの稼働する環境

EGLはNECのN88-BASIC(86) (MS-DOS Version) で開発されたソースコードで約5000命令程のプログラムであり、コンパイラによってPコードに変換されている。BASICの利用できる主記憶の関係で、語形編集用プログラムと地図

作成用プログラムとの二つのファイルに分かれてディスクに格納されているが、メニューによって相互に呼び出すことが可能である。(その他にCで開発した数本の付属プログラムがある)。実行環境はハードウェアではPC-9801(384Kバイト以上の主記憶とフロッピーディスク2台)、640×400ドットのCRT、およびプリンターが最低限必要とされる。ソフトウェアでは、MS-DOS(2.11版以上)本体とコマンドインタプリタ(COMMAND.COM)およびPコードの実行時ライブラリファイル(N88BASIC.LIB)が必要である。

この他にデータ作成用にワープロソフト(一太郎などMS-DOSのテキストファイルを出力可能なもの)や、地図の入力用にデジタイザ(LogitecのMY PAD-A3など)があれば便利である。

3 EGLによる語形データ編集

EGLが処理するデータは下の形式にまとめられたレコードである。

地点番号	語形1	語形2	語形3	コメント
------	-----	-----	-----	------

データは調査地点単位にまとめられ、1地点で最大3個までの併用形を扱うことができる。それ以上の語形や内省はコメントに記入する。(なおデータファイルが調査語形ごとに独立して作成される点はSEALと同じである。GLAPSはこの点がより自由である。)

この形式のレコードに対し、EGLは以下の順番で編集作業を行う。異なり語の検索 → 異なり語の合併 → 併用パターンの検索 → 併用パターンの合併。

3.1 異なり語の検索と合併

まず、データファイル中のすべての異なり語が検索され、その頻度が計算され、ASCIIコード順にソートされた一覧表が画面に表示される。言語地理学のデータは一般に頻度の高い小数の語形と頻度の低い多数の語形からなるものが多い。言語地図作成の目的(資料の提示か解釈の提示か)にもよるが、孤例を代表とする頻度の低い語形は他の語形と合併(地図において同じハンコを与える)しておきたいことがある。(明かな記入ミスや誤解などもやはり合併の対象になる。)EGLではこの合併操作を画面上のカーソルを操作して代表語形と合併対象を指定することによっておこなう。一回の合併処理が終わると合併された語形の色が変わり、代表語形の頻度が調整されて再度語形の一覧表が表示される。

また、プレビューという機能が利用でき、合併処理を実行する前に任意の語形の地理的分布を最大5語まで同時にCRTで確認することができる。これは人手によるデータ処理では地点番号をもとに頭の中で空想力にたよって行うか、語形毎に部分的な地図を作成するかせねばならなかった作業であり、実際にEGLを使用するとコンピュータの威力をもっとも感じさせる機能である。

3.2 併用パターンの検索と合併

一地点で複数の回答が得られた場合、それを回答の順に連結したものが併用パターンである。EGLで地図を作成する際に利用するのはこの併用パターンであって、異なり語そのものではない。従って、多くの地点で併用がおこなわれているデータに対しては、併用パターンを単位とした検索と合併をおこなう必要がある。(すべ

ての地点で回答がひとつだけならばこの処理は不要である。) 操作法は異なり語の検索・合併と全く同一であり、プレビューも使える。

ここで、異なり語の処理から併用パターンの処理にかけてEGLが何をおこなっているかを簡単に説明しておきたい。語形1=A, 語形2=B, 語形3=Cというレコードを(A, B, C)と書くこととし、以下のデータがあったとする。

地点1 (A, B, C)
 仮データ 地点2 (A, D, B)
 地点3 (B, A, C)
 地点4 (D, A, B)

このデータに異なり語の検索を施すとA, B, C, Dの4語形が得られる。ここで語形Aを代表形として語形Bを合併すると、データは以下のように変化する。

地点1 (A, A, C)
 地点2 (A, D, A)
 地点3 (A, A, C)
 地点4 (D, A, A)

この段階で併用パターンの検索をおこなうと、EGLはまず重複をとってデータを次のように書き改める。

地点1 (A, C,)
 地点2 (A, D,)
 地点3 (A, C,)
 地点4 (D, A,)

この後に検索が実行され、その結果(A, C)・(A, D)・(D, A)の3パターンが検出される。

併用パターンの検索には2種類のオプションが指定できる。ひとつは無視語形の指定オプションであり、例えば標準語形が調査地域の全体にわたって記録されているような場合に有効である。上記の仮データの場合、Aを無視語形に指定すると、最終的な検索結果は(C)と(D)になる。最大5個までの語形を無視語形に指定することが可能である。

もうひとつのオプションは「回答順位の無視」である。この指定があるとEGLは例えば(A, B, C)・(A, C, B)・(B, C, A)・(B, A, C)などを同一のパターンと見なして検索をおこなう。やはり仮データに対して「回答順位無視」を指定して検索した結果は(A, C)・(A, D)となる。これら2種類のオプションは同時に指定することも可能である。

EGLのユーザーは以上の処理を必要なだけ繰り返して語形の編集を進めていく(併用パターンの分析の途中から異なり語の分析に戻るようなことも可能)。最終的に満足できる編集結果が得られたら、それをファイルに出力しておく。編集済みのファイルはEGLファイルと呼ばれ、もとのデータファイルとは別のファイルである。EGLファイルの先頭には併用パターン数、各パターンとその頻度が記録され、その後には地点毎のパターンが地点番号と共に記録される。

4 EGLによる言語地図の作成

4.1 地形ファイルと地点ファイル

一般的な言語地図に表される情報をEGLでは、(1)海岸線・河川・等高線等の

地形に関する情報、(2)調査地点の地理的位置、(3)各調査地点の語形、の三つに分け、独立させて管理している。それぞれの情報を格納したファイルが(1)地形ファイル、(2)地点ファイル、(3)EGLファイルである。

地形ファイル、地点ファイルは基本的には地図上の座標を通常のXY座標系によって記録しただけのものであり、地形ファイルではその座標をたどって「ひと筆書き」の要領で地図が描かれる。地形・地点ファイルの作成にデジタイザを利用するための付属プログラムも用意されている。

4. 2 ハンコの割当処理

語形編集を終えた段階のEGLファイルには、まだどの語形(併用パターン)にどのハンコを割り当てるかの情報が記入されていない。この処理のために、まず一旦出力したEGLファイルを地図作成用のプログラムに読み込む。画面に語形編集画面と似た一覧表が表示され、その下に各種のハンコの候補が示されるので、やはり対話式に操作しながら各語形にハンコを割り当てる。現在のところハンコは「円」「矩形」「線」「三角」「小記号」「その他」の6系列に分かれ、全体で90個程のハンコが登録済みである。(今後増強の予定でいる。)

4. 3 CRTに言語地図を描く

ハンコの割当が済めばすぐに言語地図を作成することができる。EGLにはMS-DOSのCONFIG.SYSのようなシステム設定用のファイルがあり、それにデフォルトの地形ファイル・地点ファイルを指定しておけば、キーを押すだけで自動的にCRT上に地図が作成される。作成された地図はプリンタへハードコピーを出すことができる。また、地図の出来ばえが気に入らなければすぐにハンコ割当に戻ってハンコを調整することができる。CRTに地図を描くのに要する時間は地形の複雑さおよびディスクのアクセススピードによって変化する。筆者が分析中の「大山・中海周辺地域言語地図(LAND)」は複雑な海岸線の情報を含む364地点の地図であるが、8インチ2Dのフロッピーディスク(CPUクロックは8MHz)を使用してほしい45秒程度である。手作業に比べればごく簡単な分布の地図で80倍以上、少し複雑な地図の場合、200倍以上のスピードだと思われるが、この45秒が待てないという人のためには地形ファイルの描く画像をそのままディスクにセーブ/ロードするオプションも用意してある。これを利用すると地形を描く時間がほとんどなくなるので、どんな地図でも15秒程で描くことができる。

5 その他の機能

以上、EGLの基本的な動作を紹介したが、この他に次のような機能がある。(4, 5はEGL外部のプログラムとして稼働する。)

- ① 二つのEGLファイルを合併する(複合言語地図)。
- ② 語形毎に分類した地点番号一覧表を作成する(手作業の補助)。
- ③ 対話型処理の操作過程を記録する(ヒストリー機能)
- ④ データファイルの簡単なエラーチェック
- ⑤ 指定された言語特徴の有無を複数のファイルにわたってチェックし、地点毎にその結果を加算する。
- ⑥ EGLからMS-DOSのコマンドを実行する。(チャイルドプロセス)

6 今後の課題

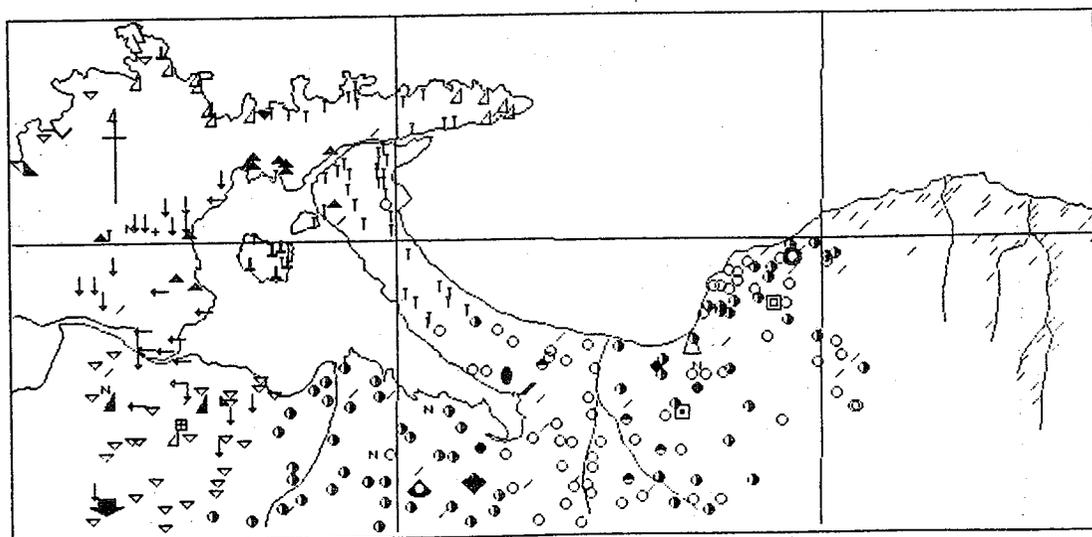
良い面ばかりを並べたててきたが、EGLはまだ問題の多い未完成のプログラムである。今後改良すべき点として少なくとも次のふたつがある。まずプロットないしはレーザービームプリンタへの出力を可能にすること。CRT上の地図はあくまで視認のためのものであり、最終的な言語地図はこれらの出力装置上に出す予定でいる。これは開発当初からの考えなのだが、まだプロット用プログラムの一部を書いたにとどまっている。次に語形の編集機能を更に充実させること。現在の編集能力は一応の基本的処理をこなせるものの、目標とした「GLAPSに近い」レベルにはまだ達していない。例えば特定の社会的属性（性別・年齢など）の条件を満たす地点だけを選択して表示するなどのことは今日の言語地理学において当然要求される機能だろうと思う。また現在、プログラムサイズが既にBASICコンパイラの限界に近づいている。全体を別の言語で書き直す必要も出てくるかもしれない。

付記

EGLはまだ開発中のプログラムですが、試用を希望される方は筆者まで連絡をいただければ実費にて配布します。その場合、試用した感想等を筆者までフィードバックしていただくことを条件とします。

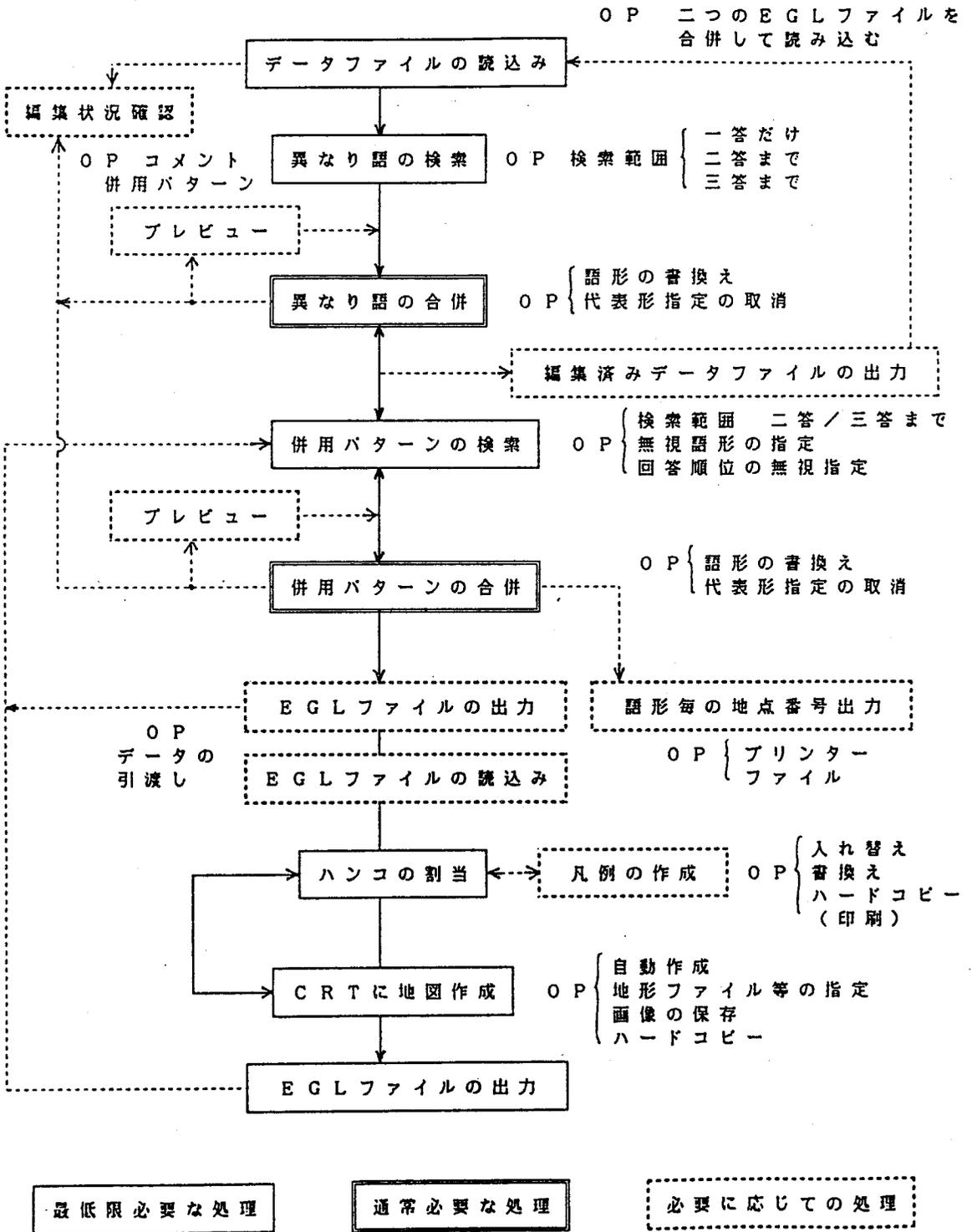
参考文献

- 荻野 綱男(1981)『方言調査分析用パッケージプログラムGLAPSの使い方』。
 私家版。
 福嶋 秩子(1983)『パソコンによる言語地理学へのアプローチ SEALユーザーズ
 マニュアル』。私家版。
 前川 喜久雄(1988)「パーソナルコンピュータ上の対話型言語地図作成支援ソフト
 ウェアEGL」, 『大山をとりまく自然環境と地域文化に関する総合的
 研究』。鳥取大学教育学部(昭和62年度教育研究学内特別経費報告書)。



付図1 EGLによる言語地図。LANDより「つらら」。5万分の1地形図境界と主要な河川を示す。

付図2 EGLによる編集と地図作成の流れ



b:ym4.dat レコト= 127

↓↑:スクロール →←:コメントON/OFF BS:併用 E:イテータ ESC:終了

46:7-220000 ホッパ`タ 林 ホ-

47:7-350000 林 ホッパ`タ[R]

48:7-470000 ホッパ`タ

Y3:年寄りを使う

49:7-510000 林 ホッパ`タ ホ-ヘ`タ

G1:普段は使わない

50:7-630000 ホ-ケ`タ ホ-ヘ`タ ホ-ヒ`タ

G1:旧, 男性が使う G3:新, 女性が使う

51:7-760000 ホッパ`タ ホケ`タ 林

52:8-130000 ホッパ`タ ホッパ`タ

53:8-200000 ホ-

54:8-380000 ホ- ホッパ`タ

Y3:おばあさんが言っているのを聞いたことがある

55:8-450000 ホッパ`タ 林

G1:普通こちらを使う

付図3 データファイルの一例。LAND より「頬」。EGLの編集状況確認画面。

b:ym4.dat 異語 第三答まで
NO 異語形

↑↓→←:スクロール F8:プレビュー ESC:終了

NO	異語形	FREQ	NO	異語形	FREQ
1	NR	(1)	2	ホ-ケ`タ	(1)
3	ホッパ`タ	(2)	4	ホッパ`	(19)
5	ホッパ`[R]		6	ホッパ`タ	(90)
7	ホッパ`タ[R]		8	ホ-	(23)
9	ホ-ケ`タ	(7)	10	ホ-タ	(5)
11	ホ-ヒ`タ		12	ホ-ヒ`タ[R]	(1)
13	ホ-ヘ`[ヒ`]タ		14	ホ-ヘ`タ	(10)
15	ホ-ヘ`タ[R]	(1)	16	ホケ`タ	(6)
17	ホタ	(5)	18	ホタカハチ	(1)
19	ホヘ`タ	(1)	20	林	(54)
21	林[R]	(1)	22	ホッパ`タ	
23		(0)	24		(0)
25		(0)	26		(0)
27		(0)	28		(0)
29		(0)	30		(0)
31		(0)	32		(0)
33		(0)	34		(0)
35		(0)	36		(0)
37		(0)	38		(0)
39		(0)	40		(0)

合併対象を選択してリターンキーを押してください (BS:指定終了)
 [現在の代表形 ホッパ`タ]

付図4 異なり語の合併処理中の画面。「ホッパタ」を代表形に選択し、次に合併対象を選択しようとするところ。

b:ym4.dat 併用 第三答まで		F:出力 ↑ ↓ → ←:スクロール F8:プレビュー ESC:終了		
NO	併用パターン	FREQ	併用パターン	FREQ
1	NR	(1)	2 ホーゲ`タ & ホー	(1)
3	ホッピ`タ	(1)	4 ホッパ`	(7)
5	ホッパ` & ホー	(3)	6 ホッパ` & ホゲ`タ	(2)
7	ホッパ` & ホ林	(5)	8 ホッパ`タ	(26)
9	ホー	(12)	10 ホー & ホーゲ`タ	(1)
11	ホー & ホーバ`タ	(1)	12 ホー & 林	(5)
13	ホーゲ`タ	(1)	14 ホーゲ`タ & ホーバ`タ	(3)
15	ホーゲ`タ & 林	(2)	16 ホー`タ	(2)
17	ホーバ`タ	(4)	18 ホゲ`タ	(3)
19	ホゲ`タ & ホッパ`		20 ホゲ`タ & 林	(1)
21	ホタ	(1)	22 ホタ & ホッパ`	(1)
23	ホタ & ホー`タ	(1)	24 ホタ & 林	(2)
25	ホタカハ`チ & ホッパ`	(1)	26 ホバ`タ & 林 & ホッピ`タ	(1)
27	林	(34)	28 林 & ホッパ`	
29	林 & ホー		30 林 & ホーゲ`タ	
31	林 & ホー`タ	(2)	32 林 & ホーバ`タ	(2)
33	林 & ホーバ`タ[R]	(1)	34 林 & ホタ	
35		(0)	36	(0)
37		(0)	38	(0)
39		(0)	40	(0)

新しい語形(20文字以内) >> No Response

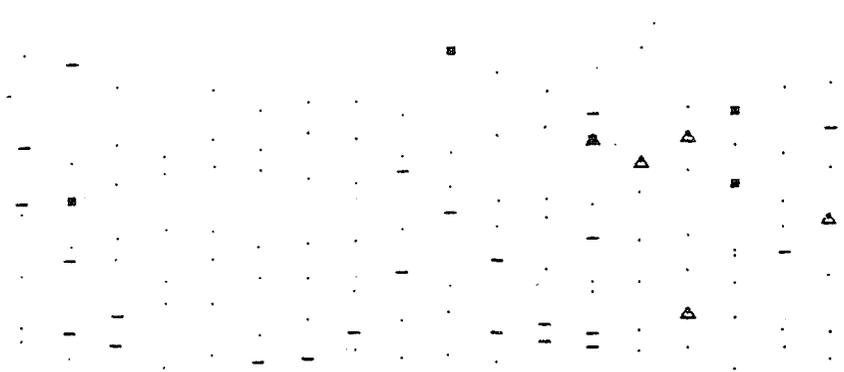
付図5 併用パターンの合併処理中の画面. 見出し語形「NR」を書き換えようとしているところ.(ホッパ`タを無視語形として、回答順位の無視も指定している.)

付図6 プレビュー画面. 「ホタ」(三角)「ホー`タ」(四角)「ホー」(線)の分布を確認している. 記号は回答順に従い青、赤、黄色で表示. (cf.付図8)

b:ym4.dat PREVIEW

回答順 1 2 3

ESC:編集



1 ○ 2 ■ 3 - 4 | 5 △
 ホー`タ の記号 >> 2

記号割当 ym4.egl

F6,7:画面↓↑

B:系列選択

NO	語形	頻度	NO	語形	頻度
1	ホッパタ	(66)	2	ホタ	(4)
3	ホー	(20)	4	ホータ	(4)
5	ホ	(6)	6	ホタ & ホータ	(1)
7	ホーパタ	(7)	8	ホータ & ホーパタ	(3)
9	ホーパ	(4)	10	ホーパタ[Reminded]	(1)
11	ホータ	(6)	12	ホタカチ	(1)
13	ホーパタ	(1)	14	No response	(1)
15	ホータ & ホーパタ	(1)	16	Miss	(1)
17		(0)	18		(0)
19		(0)	20		(0)
21		(0)	22		(0)
23		(0)	24		(0)
25		(0)	26		(0)
27		(0)	28		(0)
29		(0)	30		(0)

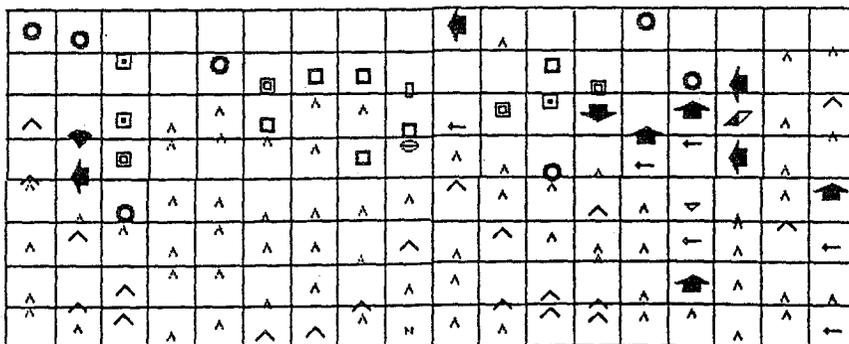
使用する記号は? >>

1 □	2 □	3 ■	4 □	5 □	6 □	7 □	8 □	9 □	10 □
11 □	12 □	13 □	14 □	15 =	16 -	17 □	18 □	19 /	20 /

付図7 ハンコ割当処理中の画面。「矩形」系列からハンコを選ぼうとするところ。

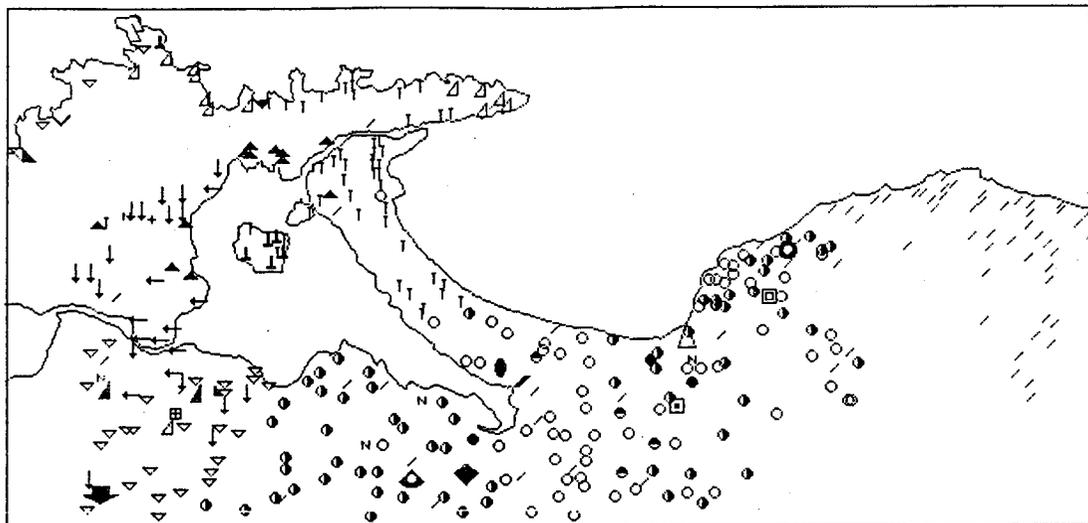
ym4.egl 地図作成

F9:メニュー K:記号 S:イメージ保存 C:コピー



4 ホー

付図8 「頬」のグロットグラムを描いたところ。



記号割当 q133.egl

F6,7:画面↓↑ A:作図 H:凡例 F:出力 C:エ°- F9:メニュー

NO	語形	頻度	NO	語形	頻度
1	ツララ	(73)	2	シンザイ	(65)
3	ツラスジ	(1)	4	シメザイ	(61)
5	サイ	(29)	6	シメダレ	(1)
7	サイゴ	(12)	8	シンジヤ	(3)
9	サイホ	(17)	10	シンザイ & カネコリ	(1)
11	サイホ	(10)	12	シンザイ & ツスラ	(1)
13	サイ & サイゴ	(1)	14	シメザイ & シンザイ	(4)
15	サイゴ & シヤゴ	(1)	16	シンザイ & サヤ	(1)
17	シヤゴ	(47)	18	サイ & シンザイ	(1)
19	シヤガ	(2)	20	シヤナI	(1)
21	シヤガ	(6)	22	シ	(1)
23	サイ & サイホ	(2)	24	カネコリ	(1)
25	サヤ	(13)	26	ゴッホ	(1)
27	コリ	(1)	28	Miss	(1)
29	シンザイホ	(1)	30	NR	(5)

付図9 LAND項目133「つらら」の地図と、ハンコ割当画面を合わせてコピーするところなる。地図の境界や河川を除いた地形ファイルを使用している。(cf.付図1)