SUMMARY

The author developed GLAPS (Generalized Linguistic Atlas Printing System) in 1975 and has since applied it, with modifications, to various field survey data. GLAPS has also been employed by other dialectologists. These applications of GLAPS reveal that this system is a useful new tool for analyzing dialect survey data even for persons ignorant of computer programming.

1. Introduction

Linguistic geography and sociolinguistics have been widely employed among dialectologists in postwar Japan. Over the last ten years, computer-processing of field survey data has become more and more common.

The author originally developed the GLAPS processor to produce linguistic atlases by computer. GLAPS has since been modified to produce glottograms and crosstables and to handle sociolinguistic data in general.

This paper presents an outline of GLAPS and an example of its usage.

2. Characteristics of GLAPS

2.1 Easy Understandability

The GLAPS processor is a FORTRAN program of about 13,000 lines. It is a package program whose strongest point is that even people ignorant of computer programming can obtain output results using it.

About thirty students of the Department of Linguistics, University of Tokyo, have used or are using GLAPS to produce crosstables from field survey data. (See, for example, Sapporo 1977 ~, 1978~.) Most of the students had never used a computer system before, but just a few hours of instruction were sufficient for them to understand how to use GLAPS and obtain their desired line-printer output.

2.2 Applicability to Various Data

GLAPS is applicable to various data, whether on fixed format cards, free format cards, or binary format disc files, and to any number of informants and variables or investigated items.

The author and University of Tokyo colleagues have applied GLAPS to data in different formats from five field surveys (Shizukuishi 1974~, 1974~, Tokunoshima 1976~, and Sapporo 1977~, 1978~). Moreover, other researchers have used GLAPS to process their own dialect data.

2.3 Compatibility with Various Computers

GLAPS is written in Japanese Industrial Standard (JIS) FORTRAN, level 7000, which is equivalent to Draft Recommendation FORTRAN of International Organization for Standardization at its maximum level (ISO Full FORTRAN) or ASA FORTRAN. It does not use assembly language and so is compatible with virtually all computer systems. In fact, GLAPS has been run on nine different computers without modification.

2.4 Flexibility with regard to Data Processing

To run GLAPS, users simply prepare their dialect data and compose a short program written in 'GLAPS language'. (There are 75 different statements in this so-called language. Some of these appear from lines 2 to 75 in Fig. 3.) In this program, the user must specify all of the functions and operations to be performed. Most programs run only 20 to 30 lines, as we shall see below.

GLAPS can perform a variety of functions needed for dialect data processing, such as the re-categorization of data, the pairing and combining of investigated word-forms, the deletion of unnecessary data, and the division of informants into subgroups by specified variables. Thus, GLAPS provides a versatile and flexible system for the user.

2.5 Processability of Multiple Answers

GLAPS resembles the SPSS (Statistical Package for the Social Sciences), originally developed at Stanford University. But GLAPS is capable of processing multiple answers often given to questions about word-form. The user simply specifies the number of answers to be accommodated in any given variable. GLAPS then automatically executes all statements related to the data and processes the specified number of answers.

3. An Example of the Application of GLAPS

As mentioned above, the author has applied GLAPS to several field studies. The following describes one of these.

3.1 Field Survey at Shizukuishi in 1974

In 1974, a team from the Department of Linguistics, University of Tokyo conducted an intensive investigation to interview all the residents of the Nishiyama area of Shizukuishi township, Iwate prefecture. The team interviewed 348 of about 500 residents above age 15, to examine distribution patterns of word-forms and the process of language change within a small area.
Fig. 1 Map of Shizukuishi

Fig. 2 Some Data Cards from Shizukuishi 1974

3.2 Data Stored in One Disc File

All the data gathered from interviews was coded and punched on 80-column data cards, and transferred onto a disc file. Fig. 2 shows some of these cards. Four data cards were prepared for each informant. The KZN cards contain information about an informant's attributes. The B10, E35, and G31 cards include answers about language usage. Though three answer fields were allowed for each language usage question, most informants gave only one or two answers to a question. Thus, on the B10 card of informant 1011, there are only twelve answers for the thirty possible answer fields.

3.3 User's Program and Output Results

Fig. 3 is a sample program, using GLAPS language, for analysis of the item 'cowlick' (the whirl of hair on the head). This figure is a fairly large program derived from many smaller programs which were used to analyze 'cowlick' trial and error. 'GLAPS' of line 1 is the top line of the program, and 'END' of line 76 indicates the end of the program. The lines starting with '*' are comment lines which the GLAPS processor ignores,
and so any useful notes or references can be entered here.

The CASES statement of line 7 denotes the number of informants, here 348. The VARIABLES statement is from line 8 to 11. If a line starts with a space (such as lines 9, 10, and 11), it means the line continues from the previous one. Names of different variables are listed in VARIABLES statements. Any words, letters, and symbols except '(', ')', '=' and ',' can be used for variable names. Unlike FORTRAN or COBOL, the length of variable names is not restricted. Lines 13 to 15 is another VARIABLES statement. But in this statement, a parenthesized three (3) follows every variable name. This means that these variables have three answer fields, that is, room for three different multiple answers to each question. The CASES and the VARIABLES (and the FORMAT of FORTRAN) are non-executable statements.

The READ statement of line 16 orders GLIPS to read ALL variables (defined by the VARIABLES statements) from input device number '7' using FORMAT statement labeled 700. Input device numbers like this are associated with data files outside a program. The number '7' here refers to the data file of Fig. 2. The FORMAT statement of lines 17 and 18 specifies data format. This is similar to the standard FORMAT statement of FORTRAN.

The TITLES statement of lines 19 and 20 gives the title of the output results, in this case, two lines in length. The title can be revised by means of a different TITLES statement if needed.

3.3.1 Map of Investigated Houses. The purpose of lines 21 to 35 is the production of a map showing the distribution pattern of the nine communities as well as informants' houses. The SIZE statement of line 24 indicates that the map size is 25 lines by 45 columns. The LOCATION statement of line 25 indicates which variables to use for location decisions. In this case, they are NORTH/SOUTH and EAST/WEST. The PRETITLES statement of line 26 and the POSTTITLES statement of line 27 indicate the character strings to be printed at the top and the bottom of the map, respectively. The DELETE statement of line 28 deletes informants with informant-numbers from 2 to 9, that is, it selects the first informant from each family.

The NAMES statement of lines 29 to 32 identifies the meaning of numbers used in the coded data. For example, the code number l of line 29 indicates 'Tate', and so on. The SYMBOLS statement of line 33 assigns symbols (including numbers, as in this case) for the numbers of the data code, for the purpose of mapping the data. This allows for much greater flexibility of design. The ATLAS statement of line 34 is an

---

**Fig. 4**

---

**Fig. 5**

---

-607-
### Age Distribution of Cases

**Table 1: Age Distribution by Community**

| Community       | Number of Cases | Over 50 | Over 30 | Over 20 | Over 10 | Total |
|-----------------|-----------------|---------|---------|---------|---------|-------|
| Tate            | 213             | 5       | 2       | 1       | 0       | 266   |
| Shinogamori     | 236             | 9       | 2       | 1       | 0       | 270   |
| Shinokawabe     | 220             | 8       | 2       | 1       | 0       | 255   |
| Hayasaka        | 220             | 9       | 2       | 1       | 0       | 255   |
| Higashi-Hayasaka| 220             | 8       | 2       | 1       | 0       | 255   |
| Kami-Shinzaki   | 210             | 7       | 2       | 1       | 0       | 248   |
| Shino-Shinzaki  | 220             | 7       | 2       | 1       | 0       | 248   |
| Higashi-Shinzaki| 220             | 8       | 2       | 1       | 0       | 255   |
| Nish-Shinzaki   | 220             | 10      | 2       | 1       | 0       | 255   |

**Note:** The table shows the number of cases over different age groups for each community. The total number of cases is 348.

### Rough Classification

**Table 2: Classification by Age Group**

| Age Group | Number of Cases |
|-----------|-----------------|
| Over 50   | 19               |
| Over 30   | 68               |
| Over 20   | 135              |
| Over 10   | 75               |

**Note:** The table shows the number of cases in each age group, categorized into four groups: Over 50, Over 40, Over 20, and Over 10. The total number of cases is 348.
instruction for the output of a map.

Fig. 4 is the output result of this ATLAS statement. This figure shows the distribution of the nine communities, plotting the locations of all informants' houses.

The NDELETE statement of line 35 of Fig. 3 cancels the effect of the DELETE statement of line 28, that is, the GLAPS processor begins to treat all the informants hereafter.

### 3.3.2 Crosstables

Lines 36 to 54 are for the production of crosstables. This is the first of three steps in our analysis of 'cowlick'. The SUBTITLES statement of lines 39 and 40 gives a more detailed explanation of the meaning of a variable — in this case, CWOLICK. The IGNORE statement of line 41 orders that those data codes for CWOLICK indicated on this line be ignored.

The RECODE statement of lines 42 to 44 is for re-categorization. In the original dialect data, informants' answers were coded separately from other variants. But by using this RECODE statement, a new code is substituted for the original and a variety of codes put together. The NAMES statement of lines 45 to 47 associates the new code numbers with specific word-forms. The CONTROL statement of line 59 instructs GLAPS to divide informants into subgroups by AGE and to print out maps for every age group.

Since AGE was recoded into seven categories on lines 48 to 51, seven maps of CWOLICK—Figs. 9 to 15—are produced by the single ATLAS statement of line 60.

Fig. 9 is for persons over 70, Fig. 10 for persons over 60, and so on. Fig. 9 shows a clear contrast between east and west. The eastern part uses 'makizyumonzi' whereas the western part 'makiguri'. These seven maps show a great difference between east and west. This suggests that glottograms of both sides of the

### 3.3.3 Linguistic Maps

There are two ways to examine the combined influence of age and geography on the word 'cowlick': by linguistic maps and by glottograms. Lines 55 to 61 of Fig. 3 are instructions for producing linguistic maps classified by age.

The CONTROL statement of line 59 instructs GLAPS to divide informants into subgroups by AGE and to print out maps for every age group. Since AGE was recoded into seven categories on lines 48 to 51, seven maps of CWOLICK—Figs. 9 to 15—are produced by the single ATLAS statement of line 60.

Fig. 9 is for persons over 70, Fig. 10 for persons over 60, and so on. Fig. 9 shows a clear contrast between east and west. The eastern part uses 'makizyumonzi' whereas the western part 'makiguri'. These seven maps show a great difference between east and west. This suggests that glottograms of both sides of the

---

Fig. 8
river might be revealing.

The CONTROL statement of line 61 of Fig. 3 erases the effect of the CONTROL AGE statement of line 59.

3.3.4 Glottograms. Lines 62 to 75 of Fig. 3 contain instructions for producing glottograms. The SIZE statement of line 65 changes the map size to 25 lines by 64 columns. The RESTORE statement of line 66 orders the restoration of original data, in this case AGE data. Hereafter, codes for AGE are not from 1 to 7 but are five digits as in the original.

The LOCATION statement of line 67 employs AGE and geography to make glottograms. The RECODE, NAMES, and CONTROL statements of lines 70 to 72 divide the investigated area into east and west. The ATLAS statement of line 73 orders two glottograms for the variable COWLICK, shown in Figs. 16 and 17. Since the CONTROL statement of line 74 specifies two variables, the ATLAS statement of line 75 produces four glottograms, that is, Figs. 18 to 21.

Figs. 18 and 19 plot the data of Fig. 16 (east side) according to whether informants are native or non-native, respectively. In Fig. 18, 'maruhosi' (symbol 0) is used by only younger informants. However, in Fig. 19, it is widely used by all generations. This means that 'maruhosi' was brought into the east by non-native speakers and that young native speakers only recently began to use it. 'Maruhosi' is thus a fairly new word-form in this area.
meaning that 'maruhosi' took root earlier in the west than in the east. Even older native informants use 'maruhosi', even among natives and non-natives. Both 'uzumaki' and 'maruhosi' are widely used in the town of Shizukuishi (Shizukuishi 1973). Past research has shown that new figures must be omitted here, Figs. 18 to 21 suggest the changes in terminology used for 'cowlick' in this area as Fig. 22.

3.4 Interpretation of the Results
4. Conclusion

GLAPS is a convenient system easily accessible to dialectologists. Moreover, GLAPS may help create a new field of 'sociolinguistic geography'. Including sociolinguistic variables in linguistic geography research will enable us to gain a more sophisticated understanding of dialect distribution patterns.

In the past, dialectologists made no use of computer facilities. Recent dialect research teams, however, especially those involved with sociolinguistic field surveys, have found computers to be useful and efficient. GLAPS is meant as an aid for researchers who are professionals in field linguistics but amateurs in computer programming.

In the humanities, generally, a package program like GLAPS could play an important role. Japan, at least, is backward in training persons in the humanities in computer programming. As far as the author knows, Japan is also backward in the development of convenient program packages for humanists. GLAPS might help promote the spread of computational dialectology, especially computer-assisted dialectology. Needless to say, equipping students of the humanities with computer facilities is most necessary.

NOTES

1Shizukuishi 1973: A survey of the linguistic geography of Shizukuishi township, Iwate prefecture, unpublished.
2Shizukuishi 1974: An intensive investigation of the Nishiyama area of Shizukuishi town, described in section 3 of this paper, unpublished.
3Tokunoshima 1976: A linguistics geography of the small island of Tokunoshima, Kagoshima prefecture. See Sibata, Takesi, et al. (eds) The Language of Amami-Tokunoshima (Tokyo: Akiyama shoten, 1977), in Japanese.
4Sapporo 1977: Sociolinguistic research on honorific expressions of Sapporo, Hokkaido. See Sibata, Takesi (ed.) Urbanization and Honorific Expressions: Sapporo 1977 (Tokyo: Dept. of Linguistics, Univ. of Tokyo, 1979), in Japanese.
5Sapporo 1978: Sociolinguistic research on honorific expressions of Sapporo, Hokkaido. See Ogino, Tsunao, et al. Sociolinguistic Study of Honorific Expressions in a Japanese City: Sapporo 1978 (Tokyo: Dept. of Linguistics, Univ. of Tokyo, 1980), in Japanese.
6See, for example, Inoue, Fumio, et al. Atlas of New Dialects in Mōgami District (Tokyo: Tokyo Univ. of Foreign Studies, 1980), in Japanese.
