בשא אד א אדשיבניפנ

Proceedings of the First Ashurbanipal Library Computer Conference

Topic

Assyrian Word Processing

May 27, May 28 1989, 6739

Peter (Jasim) BetBasoo Editor, Conference Chairman

محجحها فعسموهم فاعموهم عمدمه جدر کمجه

The Ashurbanipal Library Press Chicago August 1989, 6739

A

Proposed

Syriac

Computer

Keyboard

Layout

George Kiraz Alaph Beth Computer Systems

A PROPOSED SYRIAC COMPUTER KEYBOARD LAYOUT

By George Anton Kiraz Alaph Beth Computer Systems

With the rapid development of personal computers and their wide usage among scholars and common users, a few Syriac software programs (mainly word processors) have started to appear. These software programs are mainly designed to work with IBM-type personal computers and the Macintosh. The programs are designed by companies or individuals who saw that a need for Syriac software exists. Obviously, each programer or designer needed a keyboard layout to implement the Syriac characters on the computer keyboard. As a result, a number of keyboard layouts appeared, each of them different from the other. Since the development of Syriac software is at its early stages, it would be wise to adapt a standard keyboard layout. It will be very difficult in the future to adapt one when a substantial amount of different layouts will be available.

Since Syriac is not an international language and its usage is rather limited, the task of standardizing a Syriac keyboard layout becomes more difficult. After all those who are going to use Syriac computer software are going to be either natives of the Syriac language and/or scholars. Some might be using it on a regular basis, while others on a limited basis. Some might prefer a phonetic layout, while others a layout based on the frequency of character occurrence. Some might be using another layout of a related Semitic language (say Hebrew or Arabic) and would rather have a similar Syriac layout. An additional problem is that some programmers or designers might implement contextual analysis in the software, where the program will automatically choose the shape of characters (stand alone, initial, middle or final); on the other hand, other programers might use an ending common shape for all characters, while others might use all forms entered manually by the user. Our solution shall minimize this problem. Before proposing a Syriac keyboard layout, the following factors should be taken into consideration:

- The frequency of occurrence of characters.
- The frequency of characters with respect to other characters.
- The availability of contextual analysis.
- Placing characters to their corresponding phonetic keys.
- Minimizing the usage of shift, control and alternative keys.
- Taking into consideration other existing Semitic layouts such as Hebrew and Arabic.

From these factors we determine that it is impossible to propose one layout to satisfy all the conditions; therefore, this proposal shall include two layouts: The first we shall call the Syriac Standard Keyboard Layout (SSKL), and the second the Syriac Phonetic Keyboard Layout (SPKL).

For the SSKL, the priorities of assigning characters to the keyboard keys are as follows:

- The frequency of occurrence of characters.
- The frequency of characters with respect to other characters.
- Minimizing the usage of shift, control and alternative keys.
- Taking into consideration other existing Semitic layouts such as Hebrew and Arabic.
- Placing characters to their corresponding phonetic keys.

For example, a character whose occurrence frequency is very high, such as *Alaph*, will be assigned a key in the center of the keyboard and not on the 'a' key. However, if the assignment of a character according to its occurrence frequency is close to another key which has its phonetic value, then the phonetic key will be assigned. For example, if *Semkath* has an occurrence frequency that places it on the 'd' key, it will be shifted one position to the left to be assigned on the 's' key.

For the SPKL, the priorities of assigning characters are as follows:

- Placing characters to their corresponding phonetic keys.
- Considering other existing Semitic layouts.
- The frequency of occurrence of characters.
- The frequency of characters with respect to other characters.
- Minimizing the usage of the shift, control and alternative keys.

For example, the letter *Alaph* will be assigned to the 'a' key, even if its occurrence frequency is the highest.

To study the character frequency, a program was designed to read Syriac electronic text files. The files included all the Syriac New Testament, a total of 637,164 bytes (464,615 Syriac characters). Two tables are listed in the appendix. The first gives the character frequency of Syriac letters in the entire New Testament. The second gives the frequency of Syriac letters with respect to other letters. Bar diagrams illustrate the character frequency.

From the first table, we notice that the vowel letters *Alaph*, *Waw* and *Yudh* have the highest frequencies. Out of the total number of characters in the files (464,615 characters), *Alaph* occurred 64536 times, *Waw* 47113 and *Yudh* 41762. As for the consonant letters, *Nun* has the

highest frequency (44594 times), followed by *Lamadh* (34311 times). All other letters have frequencies lower than 30000. *Sadhe* has the lowest frequency (1514 times).

From the second table, we notice that the highest ten frequencies of letters with respect to other letters are as follows: Yudh-Nun 10193, Waw-Nun 9851, He-Waw 8890, Lamadh-Alaph 6660, Taw-Alaph 6249, Nun-Alaph 6130, Waw-Alaph 5425, Alaph-Nun 5341, Lamadh-He 5184 and Mim-Nun 4998. If we take a close look at these figures, we recognize that the noticeable high frequency of Yudh-Nun is affected by the high occurrence of plural personal pronouns (2nd feminine antein, 3rd feminine enein & henein), plural demonstrative pronouns (halein and hanein), the interrogative pronoun aylein, the 2nd and 3rd feminine suffixes of perfect verbs, the 2nd feminine suffix of imperfect verbs, and the objective pronominal 2nd feminine suffix ken. In a similar manner, we notice that the high frequencies of He-Waw and Waw-Nun, generating He-Waw-Nun, is affected by the high occurrence of the 3rd masculine pronominal suffix hun. Letters that constitute more than one word with different meanings, such as Mim-Nun in man and men, are also on the top ten list. In addition, the high occurrence of the 1st singular personal pronoun ana affects the frequency of Alaph-Nun and Nun-Alaph. The high frequency of Taw-Alaph is also affected by the feminine adjective suffix ta.

Keyboard diagrams, located in the appendix, present the proposed Syriac keyboards. The two keyboards assume the implementation of contextual analysis in the programs. If a certain program does not support this feature, it is recommended that the programer designs the characters in such a way that one shape can be used for initial and middle and another shape for final and stand alone. In this case, the shift keys should be used for final and stand alone characters. For example, the *Beth* in the SPKL will have its initial and middle shape on the 'b' key and the final and stand alone shape on the 'B' key. In the case were all the shapes are implemented in the program and are entered manually by the user, it is recommended that the lower case keys be used for initial shapes, the shift keys for middle, the alternative keys for final and the control keys for stand alone (the latest design is not recommended).

In some cases a programer might include two Syriac styles, say Estrangela and Eastern Syriac, in one font. The programer in this case can use the lower case characters for one style and the upper case for the other style. However, including two styles in one font is not efficient in terms of programing and such design should be avoided. It is recommended that every

software program include the three Syriac styles (Estrangela, Western Syriac and Eastern Syriac). The user should be able to switch between styles by a command (whether a key command or choosing an item from a menu). This way every character will have the same keys assignment in all styles. The *Alaph*, for example, will be assigned on the 'a' keys in the SPKL regardless of the style being used.

Vowels and diacritical marks can either be placed on the function keys or on the regular keys. If assigned to the regular keys, it is recommended that the vowels be placed on the shift keys of the corresponding phonetic value, for example *Fthaha* on shift 'a.' Another solution would be placing the vowels on five (or seven) adjacent shift characters. When assigning the vowels to the keyboard keys, it should be kept in mind that when the user inputs a fully vocalized text, the text should be entered first then the vowels should be placed. Typing text and vowels at the same time delays the speed of typing by a large factor.

As for the numbers and punctuation marks, they should be assigned to their corresponding keys on the keyboard. Traditional Syriac punctuation marks can be assigned to punctuation keys that are not used (such as $< > | \setminus \sim$ etc.).

A final word about the two keyboard layouts presented in this paper, it is recommended that this Computer Conference studies the layouts thoroughly making the necessary modifications. Standardizing Syriac Keyboard layouts cannot be achieved without the input and cooperation of all participants.

It is hoped that this study will help to adapt standard Syriac keyboards for computer software.

George A. Kiraz Alaph Beth Computer Systems P. O. Box 74628 Los Angeles, Ca 90004 U.S.A.

APPENDIX

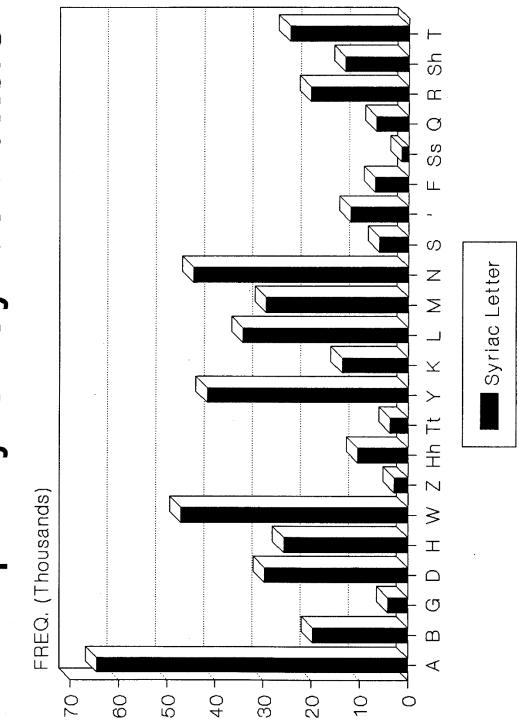
FREQUENCIES OF SYRIAC LETTERS IN THE NEW TESTAMENT:

Alaph:	64536
Beth:	19750
Gomal:	4127
Dolath:	29761
He:	25765
Waw:	47113
Zain:	2890
Heth:	10478
Teth:	3807
Yoodh:	41762
Kof:	13725
Lomadh:	34311
Mim:	29519
Nun:	44594
Simkath:	6078
'E:	11987
Fe:	6939
Sodhe:	1514
Qof:	6726
Rish:	20262
Shin:	13190
Taw:	24624
TOTAL: 4	64615

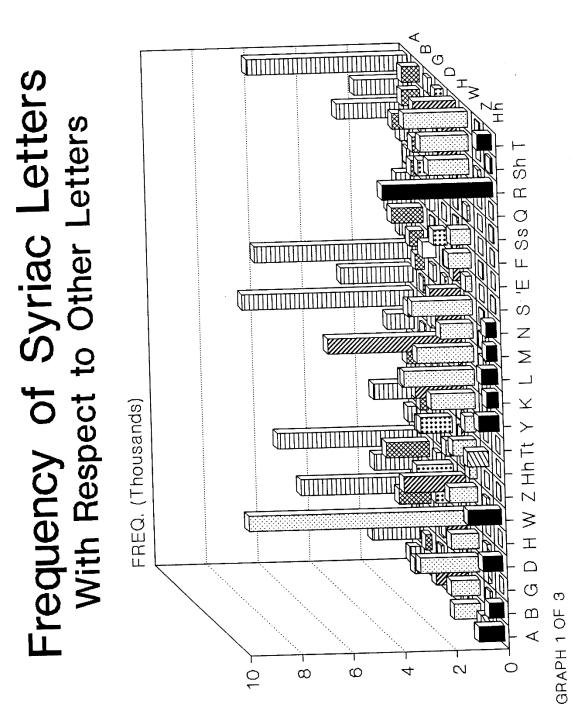
R Sh	824 601	2366 532	372 28	755 799	391 193	1499 942	1/7 0		618 275	618 275 313 27	618 275 618 275 313 27 2082 1832	618 275 618 275 313 27 2082 1832 497 59	618 275 618 275 313 27 2082 1832 497 59 93 253	618 275 618 275 313 27 2082 1832 497 59 93 253 3869 1757	618 275 618 275 313 27 2082 1832 497 59 93 253 93 253 3869 1757 104 1802	618 275 618 275 313 27 2082 1832 497 59 93 253 3869 1757 104 1802 546 0	618 275 618 275 313 27 2082 1832 497 59 93 253 93 253 3869 1757 104 1802 546 0 546 0	618 275 618 275 313 27 2082 1832 497 59 93 253 93 255 93 2555 93 2555 93 2555 93 25555 94 255555 95 25555555555555	618 275 618 275 313 27 2082 1832 497 59 93 253 93869 1757 104 1802 546 0 546 0 546 0 546 0 546 0 546 0 546 0	 618 618 275 618 275 2082 1832 497 59 93 253 3869 1757 93 262 7 262 7 262 7 657 499 94 0 94 0 950 127 	618 275 618 275 313 27 2082 1832 497 59 93 253 3869 1757 104 1802 546 0 546 0 262 7 657 499 94 0 950 127 144 446	618 275 436 618 275 436 313 27 35 2082 1832 491 93 253 1153 93 253 1153 93 253 1153 93 253 1153 93 253 1153 93 254 0 104 1802 2495 546 0 201 657 499 534 950 127 144 950 127 144 950 127 144 950 127 144 950 127 146 950 127 146 950 127 146 950 127 146 950 127 146
ð	132	389	0	590	0	854	82		62	62 3	62 3 449	62 3 449 0	62 3 449 0 294	62 3 449 0 294 265	62 3 449 0 294 265 294	62 3 449 0 294 265 294 293	62 3 449 0 294 265 294 293 93	62 3 449 0 294 265 265 294 294 33 93 580	62 3 449 0 294 265 294 293 93 93 148 148 780 0	62 3 449 0 294 265 294 265 294 294 148 1	62 3 449 0 294 265 294 93 93 93 93 93 93 93 93 148 148 1148 1190	62 3 449 0 294 265 294 280 93 148 148 1190 1190 209
																						63 55 0 126 92 92 92 92 92 92 92 92 92 92 92 92 92
1																						60 60 1177 1177 1177 1177 1177 1177 1177
																						0 143 179 541 960 960 960 185 72 26 64 64 64 818 818
	548		1	297	20	1216	0	0.7 1	142	142	142 7 3 253	142 7 3 253 231	144 7 3 253 231 231 119	142 7 3 253 231 231 119 414	142 7 3 253 231 231 119 414 414 475	142 7 7 231 231 231 119 414 414 415 415 25	142 7 7 231 231 119 414 475 475 25 25 196	142 7 231 231 119 414 475 475 25 25 139 139	142 7 7 231 231 119 414 414 415 415 475 25 25 139 139 139	142 7 7 231 231 119 414 475 475 475 25 25 139 139 67 139 67 139	142 7 7 231 119 414 475 475 475 25 25 25 139 0 0 67 130 130 139 139	142 7 7 231 119 414 475 475 475 475 475 25 139 0 139 0 139 0 180
Z	5341	1257	111	1723	1373	9851	76	527	Ì	261	261 1019	261 1019 1124	261 1019 1124 701	261 1019 1124 1124 701 4998	261 1019 1124 1124 701 4998 1045	261 261 1019 1124 701 701 4998 4998 1045 1045	261 261 1019 1124 701 701 4998 4998 1045 1045 1045 161 161 329	261 261 1019 1124 10124 4998 4998 4998 4998 1045 1045 1045 329 3306	261 261 1019 1124 10124 201 1045 1045 1045 329 306 306	261 1019 1124 1124 4998 4998 4998 4998 4998 1045 1045 329 329 329 3255 255	261 261 1019 701 701 4998 4998 4998 4998 1045 306 306 306 72 72 72 72 1048	261 261 10193 1124 701 4998 4998 1045 329 329 329 329 329 329 329 329 329 329
Σi		527																				39 1656 155 155 2860 436 436 2860 7 7 194 194 194 194 1477
	2798	745	520	1422	690	2383	464	447		1092	1092 2074	1092 2074 2184	1092 2074 2184 482	1092 2074 2184 2184 482 482 1741	1092 2074 2184 2184 482 1741 1741 36	1092 2074 2184 2184 482 482 1741 1741 36 36	1092 2074 2184 482 482 1741 137 137 137 2390	1092 2074 2184 2184 482 482 1741 1741 137 2390 2390 2390	1092 2074 2184 482 482 482 1741 137 137 2390 472 472	1092 2074 2184 482 482 482 482 137 137 137 137 2390 2390 2390 2390 472 472 419	1092 2074 2184 482 482 482 36 1741 137 137 2390 472 472 419 419	1092 2074 2184 482 482 1741 137 137 137 2390 472 182 472 182 419 22 22 23
×i	555	678	0	683	563	1099	63	126		11	$\frac{11}{1896}$	11 1896 43	11 1896 43 2002	11 1896 43 43 2002 562	11 1896 43 43 2002 562 562 389	11 1896 43 43 2002 562 562 389 389 189	11 1896 43 43 2002 562 562 389 389 189 189 29	11 1896 43 43 2002 562 562 389 389 189 189 29 29 29 122	11 1896 43 43 2002 562 562 389 389 189 189 29 29 29 29 29 29 29 29 20 20 20 20 20 20 20 20 20 20 20 20 20	11 1896 43 43 562 562 389 389 389 189 29 29 29 29 29 29 29 29 29 29 29 29 29	11 1896 43 43 2002 562 562 389 389 189 189 189 122 122 122 122 122 470	11 1896 43 43 562 562 389 189 29 29 29 29 122 122 122 29 470 553
Y	4197	2510	1675	4712	3633	1255	435	1316		636	636 68	636 68 974	636 68 974 4229	636 68 974 4229 2277	636 68 974 4229 4229 2277 2027	636 68 974 4229 4229 2277 2027 783	636 68 974 4229 4229 2277 2027 783 783 783	636 68 974 4229 4229 2277 2027 783 783 783 783 670	636 68 974 4229 4229 2277 783 783 783 783 783 1066 670 670	636 68 974 4229 4229 2277 2027 783 783 783 783 161 161 161	636 68 974 4229 4229 2277 783 783 783 783 783 670 670 161 161 163 3068	636 68 974 4229 4227 2277 783 783 783 783 783 670 670 670 633 3068 1153
It	24	175	0	111	44	168	4	321		18	18 126	18 126 0	18 126 0 254	18 126 0 254 946	18 126 0 254 946 210	18 126 0 254 946 946 210 194	18 126 0 254 946 210 210 194 31	18 126 0 254 946 946 210 1194 31 49	18 126 0 254 946 946 194 194 31 28 28	18 126 0 254 946 946 194 194 210 49 28 28 28 28 193	18 126 0 254 946 210 210 31 49 49 28 28 28 193 109	18 1266 0 9466 9466 31 194 49 49 49 49 49 40 1109 48 48 48 48 48 48 48 48 48 48 48 48 48
Æ	1014	591	9	790	0	1303	25	0		ŝ	3 804	3 804 475	3 804 475 659	3 804 475 659 422	3 804 475 659 422 441	3 804 475 475 659 659 441 441 37	3 804 475 475 659 422 441 441 37 37 0	3 804 475 475 475 422 441 441 37 37 25	3 804 475 475 475 422 441 37 37 37 25 25 47	3 804 475 475 659 659 422 441 37 37 25 25 27 27	3 804 475 475 422 422 441 37 37 25 27 47 47 429	3 804 475 475 441 37 441 37 47 47 47 47 47 47 89 88
2	453	91	134	157	, i	123	ŝ	813		0	30	0 30 2	0 30 2 107	0 30 2 107 99	0 30 2 107 99 43	0 30 2 2 99 99 43 0	0 30 2 2 99 99 43 6 0 37	0 30 2 99 99 43 43 0 37 0	0 30 2 99 99 43 43 0 37 0 0	0 30 2 99 99 43 43 6 37 0 0 0 0	0 30 2 99 99 43 43 0 0 0 0 0 0 0 128	0 30 2 99 99 43 6 0 37 0 0 0 0 128 0 0 0 0 0
м	42 1016	1119	228	824 1075	8890	2384 1100	300	920		474	424 1713	424 1713 2719	424 1713 2719 2194	424 1713 2719 2194 1266	106 424 1846 1713 160 2719 5184 2194 993 1266 1182 2344	424 1713 2719 2194 1266 2344 2344 260	424 1713 2719 2194 1266 2344 2344 893	424 1713 2719 2719 2194 1266 2344 260 893 893	424 1713 2719 27194 1266 2344 260 893 893 775 118	424 1713 2719 27194 2194 1266 2344 2344 893 893 893 893 118 118	424 1713 2719 2194 1266 2344 260 893 893 893 775 118 118 1583	11 8 106 424 9 63 1890 1846 1713 1 0 1250 160 2719 2 242 305 5184 2194 0 95 1071 993 1266 7 129 137 1182 2344 0 536 40 253 260 2 45 546 97 893 0 197 8 69 775 8 0 32 27 118 4 0 1058 34 964 1 294 171 753 1583 5 28 269 555 1874
H	42	1236 1119	33		20	2384	34	104	106	2	1846	1846 160	1846 160 5184	160 160 5184 993	1846 1846 160 5184 993 1182	1846 1846 160 5184 993 1182 253	1846 160 5184 993 1182 253 97	1846 160 5184 993 993 1182 253 253 97 69	1846 160 5184 993 993 253 253 97 97 69 27	1846 160 5184 993 993 1182 253 253 253 97 97 97 34	1846 160 5184 993 97 97 69 253 253 34 753	160 160 5184 993 97 97 97 97 97 97 34 753 753
0	165	1364	167	307	724	1425	281	1228	8	J	1890	1890 1250	1890 1250 305	1890 1250 305 1071	1890 1250 305 1071 137	1890 1250 305 1071 137 40	1890 1250 305 1071 137 40 546	1890 1250 305 305 1071 137 40 546 8	1890 1250 305 137 137 137 40 546 8 8	1890 1250 305 305 1071 137 40 546 8 8 32 32 1058	1890 1250 305 305 1071 137 40 546 8 8 32 32 171	1890 1250 305 1071 137 40 546 8 8 32 1058 171 269
U			4				15	15	11		63	63 0	63 0 242	63 0 242 95	63 0 242 95 129	63 0 242 95 129 536	63 0 242 95 129 129 536 45	63 0 242 95 95 129 536 45 197	63 0 242 95 95 129 536 45 45 197 0	63 0 242 95 129 536 45 45 197 0 0	63 0 242 95 95 129 197 0 0 294 294	63 0 242 95 95 129 197 0 0 284 28
B	695	246			500	1668	383	306	266		789	789 81	789 81 702	789 81 702 70	789 81 702 70 357	789 81 702 70 357 550	789 81 702 70 357 550 1232	789 81 702 702 357 550 550 1232 0	789 81 702 702 703 70 70 550 550 1232 0 438	789 81 702 702 70 70 70 70 550 1232 1232 0 438 438	789 81 702 702 703 70 550 1232 0 438 438 438 1101	789 81 702 702 700 700 550 1232 0 0 864 438 364 1101
Α	412	1851	111	4574	1703	5425 1	340	1649	173		3376	3376 1052	3376 1052 6660	3376 1052 6660 2796								3376 1052 6660 6130 6130 478 894 894 894 894 118 118 380 380 2799 2093
	Alaph	Beth	Gamal	Dalath	He	Waw	Zain	Heth	Teth	11 11	Indn	Yudn Kaph	rudn Kaph Lamadh	Yudh Kaph Lamadh Mim	ruan Kaph Lamadh Mim Nun	rudn Kaph Lamadh Mim Nun Semkath	rudn Kaph Lamadh Mim Nun Semkath 'E	rudn Kaph Lamadh Mim Nun Semkath 'E	rudn Kaph Lamadh Mim Nun Semkath 'E Fe Sadhe	rudn Kaph Lamadh Mim Nun Semkath 'E Fe Sadhe Qaph	rudn Kaph Lamadh Mim Nun Semkath 'E Fe Sadhe Qaph Resh	Yudh Kaph Lamadh Mim Nun Semkath 'E Fe Sadhe Qaph Resh Resh

FREQUENCY OF OCCURRANCE OF SYRIAC LETTERS WITH RESPECT TO OTHER LETTERS

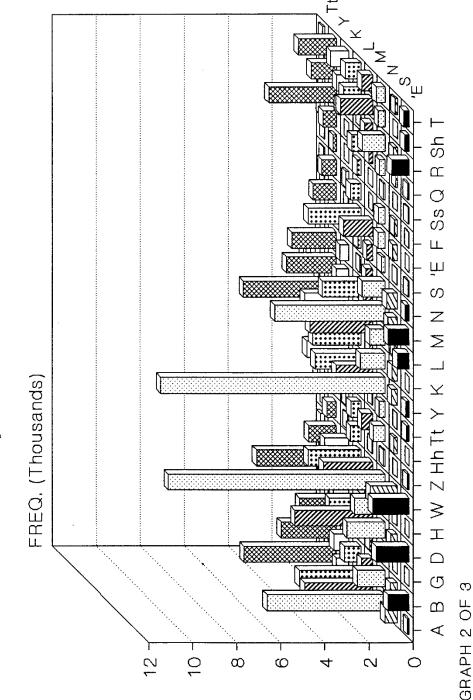
Frequency of Syriac Letters



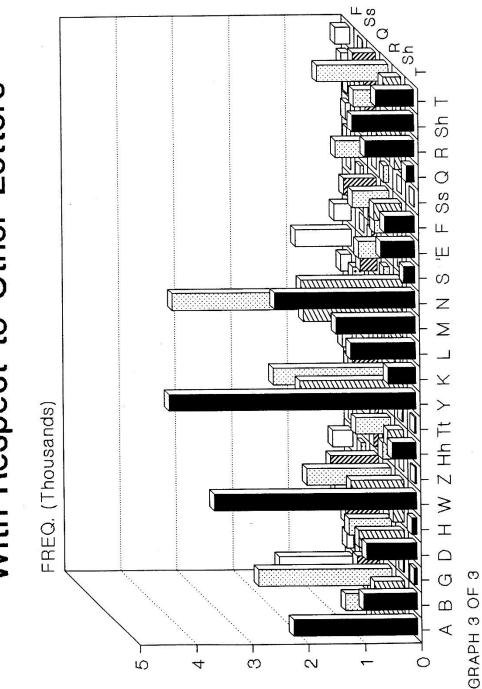
A Proposed Syriac Computer Keyboard Layout George Kiraz

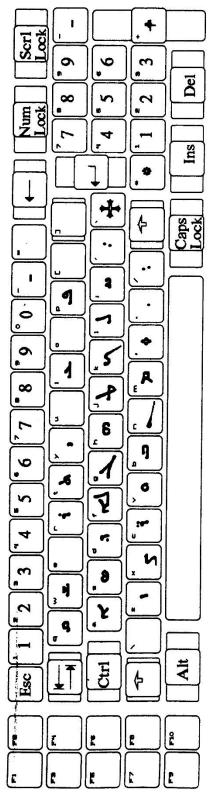


Frequency of Syriac Letters With Respect to Other Letters

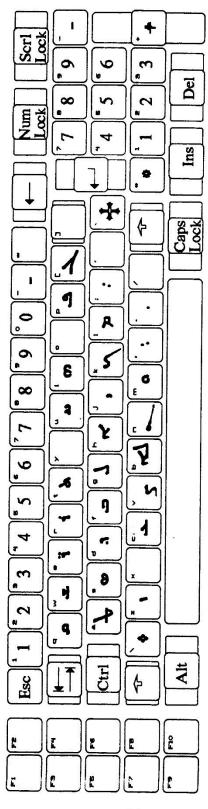


Frequency of Syriac Letters With Respect to Other Letters





PHONETIC



TANDARD