Notes and Documents

The Original Soundex Instructions

Contributed by Tony Burroughs

Census records created by the United States government are among the most heavily used records for historical research—an essential resource for genealogists, biographers, and historians alike. New users find out quickly that convenient access to the enumerations from 1880 forward requires familiarity with an esoteric, phonetic type of index called the *Soundex* that converts surnames to a combination alphanumeric code. Rules for coding surnames are found in several publications issued by the U.S. National Archives and Records Administration (NARA)[1] and in many privately published books and guides that borrow from the NARA instructions.

PROBLEMS AND VARIANCES

The Ashcroft Problem

Over the years, genealogists have discovered that the "standard" rules miscode certain names, and some researchers have discovered instructions that differ from the "standard." Roger Moffatt opened the subject for discussion in January 1997 under the topic "Specific Rules of Soundex Coding" on the Internet newsgroup <soc.genealogy.computing>. While writing a program to Soundex surnames, Moffatt had problems coding the name *Ashcroft*. Comparing various published guidelines, he found all but one to be consistent with NARA instructions.[2] The alternate instructions used a different rule for letters that carry the same number,

1. Particularly see the leaflet *Census Soundex* (Washington: National Archives and Records Administration [NARA], 1998); and the more-detailed pamphlet *Using the Census Soundex*, General Information Leaflet (GIL) 55 (Washington: NARA, 1997). NARA's guidelines also appear in *Federal Population Censuses, 1790–1890: A Catalog of National Archives Microfilm Copies of the Schedules* (Washington: National Archives Trust Fund Board [NATFB], 2001), viii–xiv. Also Anne Bruner Eales and Robert M. Kvasnicka, *Guide to Genealogical Research in the National Archives of the United States*, 3d ed. (Washington: NATFB, 2000), 18, 22; and NARA's microfilm catalogs for the 1900, 1910 and 1920 censuses.

2-5. <Rogerkiwi@aol.com> to <soc.genealogy.computing>, 21 January 1997. The exception he reported is said to be at <http://home.utah-inter.net/kinsearch/Soundex.html>, but the site could not be located on 12 November 2001. However, the software code is found online at <www.cpan.org/authors/id/CNANDOR/Text-Soundex-2.20-bin-1-Mac.readme> and <http://j-walk.com/ss/excel/tips/tip77.htm>.

when separated by *h* or *w*—a variation that generates a different Soundex code. Finding that he could locate *Ashcroft* in NARA's Soundexed records using the "different" rule, Moffatt asked if anyone knew of the whereabouts of "definitive" Soundex instructions. In response, and in a renewed discussion ("Soundex Problems") in early 1998, other programmers noted the following:

- Soundex instructions provided by the Clayton Library in Houston, Texas, also carry the different rule,[3] although the library's handout carries no reference and the staff does not recall the origin of the guide.[4]
- A 1983 computer program written by George H. Pardue generated Soundex codes according to both the NARA and Clayton Library guidelines.[5]
- Some say there are "a number of variations of the Soundex algorithm."[6]

Meanwhile, programmer Roy Katschke noted other problems. Using NARA-based instructions published in *The Source,* he could not find his own surname in the census Soundexes—although he did find the name under another code (K320 rather than K322). Investigating the matter further, he found two similar names with the same problem, *Patschke* and *Demshki.* He then examined seven genealogy software programs and found that all followed guidelines published by NARA.[7]

In sum, the programmers realized that the alternative guide provided by Clayton enables researchers to locate names on Soundex microfilm that they could not find by using the "official" instructions from NARA. They were concerned about adding these rules to computer code without identifying the origin of their instructions, but they concluded that the need to find all names in Soundexed records was an overriding imperative.

The Burroughs Problem

In 1999, while writing *Black Roots*, I included a chapter on using census records and wrote instructions for Soundex coding. In selecting an example, I recalled the difficulty I encountered in the 1970s, when I first used NARA's coding instructions (see Table 1 for a sample) to find my Burroughs ancestors on microfilmed Soundex cards for the federal censuses.

---

3. Myron Howard to <soc.genealogy.computing>, 20 January 1997. Howard's response is dated before Moffatt's question because of their differing time zones.

4. Ray Cox to <soc.genealogy.computing>, 16 February 1998. The Clayton instructions are also published online <www.hpl.lib.tx.us/clayton/cla_c1a.html>. *Note:* In the phrase "cla_c1a" within the library's URL, the middle character before the underscore is the digit 1, while the middle character after the underscore is the letter *l.*

5. Jim Elbrecht to <soc.genealogy.computing>, 23 January 1997 and 13 February 1998. In the latter message, Elbrecht referenced the computer code for the program that appeared in *Quinsept User Group Newsletter* 3 (June–July 1986): 5.

6. <Pzavon@delphi.com> to <soc.genealogy.computing>, 27 January 1997; in reporting the "number of variations," the writer added his belief that "no one is more correct than another."

7. Roy Katschke to <soc.genealogy.computing>, 12 February 1998. For the instructions first used by Katschke, see Johni Cerny and Arlene Eakle, *The Source: A Guidebook of American Genealogy* (Salt Lake City: Ancestry, 1984), 99–100. The revised edition, edited by Loretto Dennis Szucs and Sandra Hargreaves Luebking in 1997 (pp. 124–25) also uses Soundex instructions based on NARA publications.

Table 1

NARA Coding Instructions for Soundexed Names

Number	Represents Letter
1	b, f, p, v
2	c, g, j, k, q, s, x, z
3	d, t
4	l
5	m, n
6	r

To create the Soundex code, use the following steps:

Step 1: Bring down the first letter of the last name.

Step 2: Disregarding the first letter, omit the remaining letters A, E, I, O, U, W, Y, and H.

Step 3: Write down the numbers found on the Soundex Coding Guide for the first three remaining unslashed letters. Add zeros for any empty spaces. Disregard any additional letters.

ADDITIONAL RULES

Names with double letters: If the surname has any double letters, they should be treated as one letter. For example, the two "t"s in Ritter are coded with one "3" (R360).

Names with letters side-by-side that have the same number on the Soundex coding system: A surname may have different letters side-by-side that have the same number on the Soundex Coding Guide. For example, PF in Pfister (1 is the number for both P and F [P236]); CKS in Jackson (2 is the number for C, K, and S [J250]). These letters should be treated as one letter.

Names with prefixes: Mc and Mac are not considered prefixes. If the surname has a prefix, such as Van, Con, De, Di, La, or Le, code the name both with and without the prefix because it might be listed under either code.

Mixed codes: If several surnames have the same code, the cards for them are arranged alphabetically by given name. There are divider cards showing most code numbers, but not all. For instance, one divider may be numbered 350 and the next one 400. Between the two divider cards there may be names coded 353, 350, 360, 364, 365, and 355, but instead of being in numerical order they are interfiled alphabetically by first name.

Extracted from Anne Bruner Eales and Robert M. Kvasnicka, *Guide to Genealogical Research in the National Archives of the United States*, 3d ed. (Washington: NATFB, 2000), 18, 22.

Applying the standard NARA instructions to *Burroughs* produces the following:

B	first letter of code
u	vowel, not coded
r	6
r	in a double-letter, the second is not coded
o	vowel, not coded
u	vowel, not coded
g	2
h	not coded
s	2
Result:	B622

However, Burroughs is not to be found under B622 in NARA's microfilmed Soundex

to the 1880 and later censuses.[8] Hearing that some driver's license numbers carry Soundex codes,[9] I checked mine and noted that the first four characters were B620. Using that as a code to reconsult the Soundex, I found my missing Morris Burroughs.[10]

In light of this experience, my 1999 preparation for writing the Soundex discussion in *Black Roots* included a reexamination of NARA's rules. The published instructions still generated B622 for Burroughs. The "Soundex Machine," a computer program then offered at the NARA website—which automatically generated codes for any queried name—also produced B622.[11] In writing my own discussion for *Black Roots*, I was at a loss to explain the discrepancy to my readers. Moreover, during the reexamination, I noted one curiosity: B622 was *originally* penned on the Morris Burroughs card for the 1920 Soundex, but it was crossed out and replaced with B620. That left me wondering whether the original coders also had problems.

THE SEARCH FOR ORIGINAL INSTRUCTIONS

Inquiries to NARA did not resolve the problem. A staff specialist, an archivist, and a retired archivist all assured me that no one there had ever found "original instructions" for the Soundex. In preparation for personal research in the original manuscript files of the Bureau of the Census, I studied NARA's preliminary inventory to that record group and learned that the Soundex was created as a project of the Work Projects Administration (WPA).[12] Therefore, my on-site research at NARA covered not only the Bureau files (Record Group 29) but also those of the WPA (RG 69); it eventually extended to other agencies that have also used Soundex and numerous other research institutions as well. Pre-NARA instructions do indeed exist, in various forms, along with abundant documentation of the following events and circumstances.

Soundex Origins

The phonetic indexing system called Soundex was developed by Robert C. Russell of Pennsylvania, who obtained six patents for various modifications of it between 1916 and 1924.[13] Russell then licensed rights to the system to Remington-

---

8. The NARA-generated code, B622, also appears for Burroughs in a widely used, computer-generated guide to some 500,000 surnames; see Bradley W. Steuart, ed., *The Daitch-Mokotoff Soundex Reference Guide*, rev. ed., 2 vols. (Bountiful, Utah: Precision Indexing, 1994).

9. Because the Soundex code is essentially a method of locating names in a database, motor vehicle departments in some states incorporate this system into their drivers' license numbers.

10. For example, see Morris Burroughs, B620, 1910 Soundex, NARA microfilm publication M1043, roll 44.

11. Since 1999, NARA has dropped the program from its website <www.nara.gov/genealogy/soundex/soundex.html>, although the instructions for manually computing Soundex codes still appear there.

12. Katherine H. Davidson and Charlotte M. Ashby, *Records of the Bureau of the Census*, Preliminary Inventory (PI) 161 (Washington: National Archives and Records Service, 1964), 113.

13. *Official Gazette of the United States Patent Office, 1916* (Washington: Government Printing Office, 1916), 159 (no. 1,207,220); *Official Gazette, 1918*, 49–50 (no. 1,261,167) and 998 (no. 1,283,319); *Official Gazette, 1922*, 344 (no. 1,435,663) and 334–45 (no. 1,435,664); *Official Gazette, 1923*, 496 (reissue no. 15,582); and *Official Gazette, 1924*, 834 (no. 1,481,744).

Rand, who trademarked it as the *Soundex* in 1927.[14] Then several factors coalesced to make the system immensely valuable. Widespread unemployment caused by the financial crash of 1929 prompted Congress in 1935 to launch the WPA. That agency's search for public projects that would provide jobs generated a request from the Bureau of the Census. Because birth registrations did not begin in most states until after 1900, the Bureau was overwhelmed with requests from Americans who needed proof of age but had no acceptable alternate sources (e.g., Bible records or affidavits of a doctor or midwife). WPA indexing of the censuses would help the Bureau find the needed data.

After completion of the project, a problem surfaced in the testing phase. Most names could be spelled in a variety of ways. Searchers might miss the name *Stewart*, for example, if it were spelled (and indexed) as *Steuart* or *Stuart*. So, in 1936, the Bureau sought bids from private indexing companies, then chose Remington Rand to apply Soundexing procedures to the population schedule of 1900. At that point, the census extraction cards that had been indexed alphabetically were rearranged into Soundex order; and the project was completed in 1937.[15]

Rules Used

So what rules did the Soundexers use? In 1925, the Library Bureau Division of Remington Rand published an instruction manual, *Progressive Indexing and Filing*, which was revised several times in the subsequent decade.[16] In addition to several office procedures, it provides instructions for coding names under the Soundex method—including a rule not found in the instructions issued by NARA. That set of rules, which produces different Soundex codes for many names—is the set disseminated by Houston's Clayton Library.

These were also the rules followed under the Remington Rand contract for the 1900 census. They were the rules adopted by the Bureau of the Census, by 1957, to train Bureau employees who conducted age searches for the public.[17] And they were the rules adopted by other government agencies that have applied Soundexing

---

14. Trademark registration no. 0,230,958 (filed 31 March 1927, registered 9 August 1927, by Rand Kardex Bureau), file consulted at the U.S. Patent and Trademark Office, Crystal City, Virginia. In that same year, Rand Kardex Bureau merged with Remington Typewriter to form the Remington Rand Co. See Kardex Information and Materials Management Systems, *A History of Kardex Systems, Inc*. <www.kardex.com/update/history.htm> downloaded 9 November 2001. The company first used the name *Russell-Soundex* for its system, but later dropped *Russell*.

15. Davidson and Ashby, *Records of the Bureau of the Census*, 113, incorrectly state that the 1900 Soundex was compiled between 1937 and 1939. For corrective information, see "Old-Age Census Work to be Started Today," *St. Louis Daily Globe-Democrat*, 16 September 1935, p. 7B. Also Bureau of the Census Monthly Reports for September 1936 and May 1937; Monthly Reports, Jan. 1912–June 1956 (PI 161, Series 193), RG 29; National Archives Building, Washington.

16. For the instructions used to install the Census Bureau system in 1936, see Library Bureau Division, Remington Rand, *Progressive Indexing and Filing: A Text Arranged for Courses of Various Lengths* (Buffalo, New York: Remington Rand, 1934). This book, which was written in 1925, was revised in 1928, 1931, 1939 and reprinted nineteen times by 1941.

17. This was seven years before the 1880 Soundex was transferred to NARA for public use. I found the Bureau's 1957 instructions—as transcribed in the appendix to this paper—in July 2001, while combing the Bureau's manuscript files at NARA.

to their records—the Immigration and Naturalization Service[18] and the Social Security Administration[19] being two notable examples.

THE DISCREPANCIES

A Comparison: NARA vs. Original

The "lost" Ashcroft, Burroughs, Demshki, Katschke, and Patschke names are all rooted in the same problem—the manner in which the long-standing NARA instructions code *h* and *w* when they appear in a name somewhere other than the first letter. The basic differences are these:

NARA'S TRADITIONAL RULES:
- The letters *h* and *w* are treated the same as vowels.
- Neither type is coded.
- Both types are "separators" of letters with equivalent numbers.

Therefore, letters on either side of vowels, with equivalent soundex numbers and the letters *h* and *w*, are treated as two letters and both are coded. In *Vargas*, for example, both *g* and *s* are coded.

ORIGINAL SOUNDEX RULES
- The letters *h* and *w*, like vowels, are not coded.
- Vowels are separators of letters with equivalent numbers.
- *H* and *w* are *not* considered "separators" of letters with equivalent numbers.

Therefore, when *h* and *w* link two letters that have equivalent Soundex numbers, those two letters are coded as one.

To use *Burroughs* as an example:

NARA RULES		ORIGINAL RULES	
B	first letter of code	B	first letter of code
u	vowel, not coded	u	vowel, not coded
r	6	r	6
r	double-letter, not coded	r	double-letter, not coded
o	vowel, not coded	o	vowel, not coded
u	vowel, not coded	u	vowel, not coded
g	2	g	2
h	not coded	h	not coded
s	2	s	not coded (equivalent letter to *g*)
		0	added for the mandatory third digit
Result: B622		*Result:* B620	

---

18. For example, see Works Progress Administration, "Project Proposal No. 364–4," 5 January 1940, U. S. Department of Labor, Immigration and Naturalization Service (INS), Buffalo District, which proposed "to provide an adequate permanent Index System under the Russell Soundex System for records of naturalization." See file 55938/907, INS Subject Correspondence Files, entry 9; accession no. 85-58A734; Records of the Immigration and Naturalization Service, RG 85; National Archives Building. By January 1940 the Works Progress Administration had already been replaced by the Work Projects Administration; the proposal seems to have been prepared on an outdated form.

19. U.S. Social Security Administration, *Russell Soundex Coding and Filing System: A Programmed Text* ([Washington: Government Printing Office], 1969), 54.

A Correction: The Ashcraft Rule (H and W Rule)

Responding to the problem in February 2000, NARA added to its website what it calls the "Ashcraft Rule"—borrowing from the Clayton Library guidelines.[20] (At that point, I had not yet found the instructions used by the Bureau of the Census but had found the guidelines produced by the Social Security Administration for its employees.) Unfortunately, the long-awaited revision of the *Guide to Genealogical Research in the National Archives*, which was issued later in 2000, does not include the correction. Nor does the 2001 edition of *Federal Population Censuses, 1790–1890*.

On 1 September 2001, in preparation for a discussion of this issue as my keynote address at the annual Federation of Genealogical Societies conference, I asked the NARA staff about plans for updating its publications to be consistent with the Soundex instructions now published at its website. The response: no changes were planned for its printed finding aids because the original Soundex instructions had never been located. That prompted a follow-up letter from me to John Carlin, Archivist of the United States, informing him of the location of the original Soundex instructions and requesting that all finding aides, guides, and publications which list Soundex instructions be revised to reflect the missing rule. On 26 September, Diane Dimkoff, NARA's Director of Customer Services Division, replied with welcome assurance. In part, her letter states:[21]

> In preparation for the opening of the 1930 Census in April of 2002, we are reviewing the guidelines and making necessary changes. As we review and reprint other research guides and finding aids according to our publication schedule, we will update those as well. Please be assured that your recommendation will be included.

> Constance Potter, NARA genealogy specialist, coordinates NARA's finding aids of genealogical interest. She has a copy of your letter and is incorporating your suggestions into the introduction to the 1930 census catalogue and in the revisions to the Soundex worksheets. We will be training staff and volunteers regarding these changes.

Even when this change occurs, however, researchers should keep one thought in mind: those who created the original Soundex entries may not have always applied the "H and W Rule," as in the case of Morris Burroughs in 1910. Wisdom suggests coding names both with and without the rule.

---

20. The web page <http://www.nara.gov/genealogy/soundex/soundex.html> was last updated on 16 June 2001. HTML coding reveals it was last modified on 23 July 2001 but does not state what modifications were made. HTML coding for the web page <http://www.nara.gov/genealogy/coding.html> states that the site's Soundex instructions were taken from *Using the Census Soundex*, GIL 55. Revisions by NARA archivist Claire Prechtel-Kluskens of 19 February 2000 include miscellaneous coding changes and the Ashcraft Rule.

Note that while Roger Moffett's original problem dealt with *Ashcroft*, the NARA rule spells the name *Ashcraft*. Both variants code the same.

21. Letter from Diane L. Dimkoff, Customer Services Division, Access Programs, NARA, to Tony Burroughs, 26 September 2001.

CONCLUSION

The widespread omission of one of the Soundex coding rules affects more than genealogists. Biographers, historians, and other researchers also use Soundexed materials. Attorneys and judges sometimes use Soundexes for naturalization records and censuses to close probate cases. An endless array of books, guides, handouts, and websites have published incomplete Soundex instructions; and trusting archives, historical societies, libraries, and teachers continue to disseminate the deficient rules. Millions of people today, worldwide, rely upon Soundexed records in their efforts to reconstruct America's past, and untold numbers have been stymied by one unfortunate flaw in the "standard" instructions.

This paper's explanation of the problem—along with the reproduction, below, of the original instructions issued by the Bureau of the Census—should help to expedite the dissemination of correct instructions throughout the public sector.

————

Appendix

Form No. Adm/S—474

DEPARTMENT OF COMMERCE
BUREAU OF THE CENSUS

SOUNDEX
CODING RULES*

Only the *surnames* are *coded.* The given names of each group are alphabetized in a separate operation, which is taken care of later.

Consonants

Consonants only are coded. The code is as follows:
1. b f p v
2. c g j k q s x z
3. d t
4. 1
5. m n
6. r

---

* AD/S [Administrative Services Division] file "Searching Soundex Census Records[,] Age Search (STDS [Standards], and Back-up, 1957–63)"; Management Studies and Reports (Bureau of the Census Finding Aid A1, series 356); accession no. NN3-029-098-004; Records of the Bureau of the Census, RG 29.

These instructions are transcribed here in their original form, with a few minor bows to the conventions of typography. The document is a typescript, not a publication. Because typewriters of the era did not offer italics and dashes as an option, typists used underscoring for emphasis and hyphens in lieu of dashes. On typeset pages, however, the standard (then, as now) has been italics, boldface, and em (or en) dashes under particular circumstances. Conversions are made herein, as appropriate, to improve readability. Also for clarity, this transcript follows the modern typographic convention of italicizing letters that appear as "words" within text. The "H and W Rule" is also highlighted here, although it was not specially marked in the original.

Note that code numbers do not go beyond 6. The figures 7, 8 and 9 are not used.

The letters *b, f, p,* and *v* are coded as 1, or they have a value of 1. The letters *d* and *t* are coded as 3, or they have a value of 3, etc.

Any letter having the same value as another is called an equivalent of the other. Thus *c* is an equivalent of *s, g* is an equivalent of *k, b* is an equivalent of *p,* etc.

The initial, or first letter, of a surname is always included in the code and it is always printed. This applies to all *twenty-six* letters of the alphabet without exception. If a surname begins with *H,* then that letter appears first in the code; if it begins with *0,* than that letter appears in the code.

Vowels

The vowels *a, e, i, o, u,* and *y* are never coded. They cannot be coded in any event, as they have no numerical value in the code.

A few examples of very simple surnames and their correct codes follow:

Ruibler   – R 146: *R* is the initial letter; *u* and *i* being vowels are not coded; *b* has a value of 1; *l* has a value of 4; *e* being a vowel is not coded and *r* codes as 6.

Denvyk   – D 512: *D* is the initial letter; *e* being a vowel is not coded; *n* has a value of 5; *v* has a value of 1; *y* being a vowel is not coded and *k* codes as 2.

Length of Code

Every code contains the initial letter and 3 figures—*never more and never less.* In a long name we code into it until we reach three figures and then stop.

EXAMPLE:
Palosonvict – P 425: *P* is the initial letter; vowel *a* is not coded; *l* is 4; vowel *o* is not coded; *s* is 2; vowel *o* we do not code, and *n* is 5. This gives us our three figures and we do not code farther.

Where we do not have enough letters in a name to give us three figures, we add zeros on the end of the code to bring it to three figures. We *never* have a code with one zero *between* two other figures. Such a zero is *always* added on the end of the code.

EXAMPLES:
Lears   – L 620: *L,* initial letter; *e* and *a* not coded; *r* is 6; *s* is 2 and one zero added.
Lear   – L 600: *L,* initial letter; *e* and *a* not coded; *r* is 6, and two zeros added.
Lee   – L 000: *L,* initial letter; *e* and *e* not coded; three zeros added.

Separators and Non-Separators

Doubles and equivalents, *when not separated* by a vowel or a letter of a different value, *are coded as one letter* regardless of how many equivalents may occur together.

EXAMPLE:

Spinner	– S 156	*Buckskin*	– B 250	*Hennessy*	– H 520
Rodtecks	– R 320	*Sumner*	– S 560	*Hapver*	– H 160
Campbescks	– C 512	*Lasjzkort*	– L 263		

Vowels, while not coded, serve as *separators* of equivalents, or like consonants. Equivalents *so separated* by a vowel, or a consonant of a different value, are *treated as two letters* or two groups, and we code on each side of such a separator.

EXAMPLES:

Dutaton – D 335 : The *a* serves as separator and each *t* is coded.
Buckeskin – B 225 : The *e* is a separator and we code on each side of it.
Hapyver – H 116 : The *y* is a separator and we code on each side of it.
Lasjozkort – L 226 : The first *o* is a separator and we code on each side of it.
Uskazejz – U 222 : The *a* serves as one separator and the *e* as another.
Zodrt – Z 363 : The *r* is a separator as it has a different value.

The letters *h* and *w* are *completely disregarded* except as initial letters. In any other part of a surname the letter *h* and *w* are treated as if they are not there. They are not recognized as separators of like or equivalent consonants. *This is very important.*

EXAMPLES:

Hughgill – H 240 : The *h* is not there and the double *g* codes as one letter.
Kophveaux – K 120 : The *h* is not there, and the *p* and *v* are coded as one letter.
Blaschko – B 420 : The *h* is not there, and the *s, c k* code as one letter[.]
Rochwzi – R 200 : The *h* and *w* are not there and the *c* and *z* are coded as one letter.
Achwos – A 220 : The *h* and *w* are not there, but the *o* serves as a separator, so we code both the *c* and the *s*.
Zodwt – Z 300 : The *w* is not there and the *d* and *t* code as one letter.

Equivalents of Initial Letters

Equivalents of an initial letter, when *not separated* from that initial letter by a vowel, or by a consonant of a different value, are NOT coded. The initial letter and all equivalents of it not separated from it by a vowel or by a consonant of a different value are *treated as one letter*. Such a group is not coded but is covered by the initial letter itself, which *appears in the Code*. The same is true of doubles of the initial letter. This is *very important*, as it often causes many mistakes on the part of the inexperienced coder.

EXAMPLES:

Lloyd – L 300: The *L* in the code covers the *Ll* in the name.
Pfeiffer – P 160: The *P* in the code covers the *Pf* in the name.
Schmidt – S 530: The *S* in the code covers the *Sc* in the name.
Skinner – S 560: The *S* in the code covers the *Sk* in the name.
Czar – C 600: The *C* in the code covers the *Cz* in the name.
Squires – S 620: The *S* in the code covers the *Sq* in the name.

Llewelyn　　– L 450:　The *L* in the code covers the *Ll* in the name.
Gschwind　　– G530:　The *0* in the code covers the *Gsc* in the name.
Zgombiczi　　– Z 512:　The *Z* in the code covers the *Zg* in the name.
Cjolbert　　– C416:　The *C* in the code covers the *Cj* in the name.
Pfaender　　– P 536:　The *P* in the code covers the *Pf* in the name.
Scjzowski　　– S 200:　The *S* in the code covers the *Scjz* in the name.
Schwzi　　– S 000:　The *S* in the code covers the *Schwz* in the name.
Schjort　　– S 630:　The *S* in the code covers the *Schj* in the name.
Zschiesche　　– Z 200:　The *Z* in the code covers the *Zsc* in the name.

The following names furnish good coding practice:

Campbell	– C 514	*Zychs*	– Z 200	*Czar*	– C 600	Skophver	– S 160
Campabell	– C 511	*Oszecsic*	– 0 222	*Cyzar*	– C 260	Schultz	– S 432
Blaschko	– B 420	*Pfaff*	– P 100	*Cardwt*	– C 630	Tzinberg	– T 251
Blascoko	– B 422	*Squeschk*	– S 200	*Cardet*	– C 633	Swjedrt	– S 363

Prefix and Compound Names

In prefix names we code right on into the second or third part of the name until we get our three figures in the code. We *completely ignore the break in prefix names*, and treat them as though they were one word instead of in sections. Such names are very common and we see thousands of them. They are names which are essentially one name, although the prefixes divide them into two or more parts insofar as spelling is concerned,

EXAMPLES:

Von Hoffmann	– V 515	*Santa Lucia*	– S 534
De La Rouge	– D 462	*Van Dyne*	– V 535
Ten Eyck	– T 520	*SanQuentin*	– S 525
Der Rausch	– D 620	*Le Marr*	– L 560
McKittrick	– M236	*MacGrady*	– M263
Bon Vivant	– B 511	*Bella Ria*	– B 460

The prefix *Saint*, when abbreviated *St.*, is nevertheless coded as though spelled out in all cases.

EXAMPLE:
St. Daniel　　　　– S 535:　This is coded as *Saint Daniel*.

Compound names are *two surnames combined* to make one surname. They are rare, and seldom encountered. Such names are coded just *to the end of the first part of the name*.

EXAMPLES:
Schumann Heinck, Madam – S 550:　(*Heinck* would here be treated as the given name.)
Duff-Gordon, James　　– D 100:　(*James* is here clearly indicated as the given name.)

Religious Orders

In the case of a Sister, Mother, or Brother, where it is clearly indicated that the title is religious, they are coded as though the surname were *Sister* or *Mother* or *Brother*.

EXAMPLES:
Theresa Agnes   – S 236: (Given name *Theresa Agnes*).
Finn, Anne Margaret (Sister) – S 236: (Given name *Anne Margaret*—Surname ignored).
Agnes Therese (Mother)   – M360: (Given name *Agnes Therese*).

Priests almost always have surnames and the surname is coded and handled just like all other surnames.

Indian, Alaskan and Hawaiian Names

Indian, Alaskan and Hawaiian names which are often sectioned, are coded to the first break, or the end of the first section, and the second section is used as the given name.

EXAMPLES:
Ta na sa   – T 000 : (Given name here would be *na*).
E-tom-ki   – E 000 : (Given name *tom*).

The six words – *the*, *of*, *a*, *an*, *and*, *for* – when used as English words are not coded.

EXAMPLES:
The Deer Hunter   – D 600: (Given name *Hunter*).
An Old Chieftain   – O 430: (Given name *Chieftain*).

In Chinese Names we code the first name shown and alphabetize the second.

SUMMARY

To sum up, the main points to be remembered are:

Equivalents, when not separated by a vowel or a consonant of different value, *are treated as one letter,* regardless of whether there are two or ten of them, or whether they appear in the beginning, middle or end of a surname.

When equivalents or like letters are separated by a vowel or vowels, or by a letter, or letters, of a different value, we *code on both sides of such separators*.

The letters *h* and *w* are completely disregarded except as initial letters, and beyond that point they are *treated as though they are not in a name*. This requires very careful watching and is *very important*.

Equivalents or doubles of the *initial letter*, when not separated by a vowel or a consonant of a different value, are not coded, but are covered by the initial letter, which itself appears in the code. This is also *very important* and requires very careful watching.

Real proficiency in coding is only attained by practice in coding.