

From 1855 to 1872 he worked as a topographer for the War Office in London.

In 1858 he married an English lady (Welsh perhaps?) Ada Parry.

He also worked for the company G. Phillips as a geographer; this company no longer exists.

In 1909 he was awarded an Hon. Dr. (Philosophy) by Göttingen University (for his work on Martin Behaim and his globe – 1908).

In geography he is known for his expertise on immigration patterns.

He was "born on the edge of the SW German wave into the Rhineland and witnessed the immigration trends from Ireland to the United States of America and to Britain, and from the British countryside into the towns!"

His 'Laws' on these matters became – and still are – a valued text.

His contributions to cartography were highly valued; he made maps of Africa and India.

His second field of interest was exploration; in addition to explorations in Africa and India, he worked on North American topics and Russian explorations in the Far East.

He translated Vasco de Gama's first voyage (1888).

Climatology was another field of his expertise. It is said that he "sought to establish generalisations of worldwide application".

For years, I have laboured under the impression – delusion even – that E.G. Ravenstein was some sort of imported gymnastic teacher. I now take off my hat to honour a great scholar who was First President of the GGS from 1862 until 1871, and for many years also Honorary Director of the GGS and co-founder of the NOA. Such rich scholarship is fitting for an early Olympian. His contributions to Olympism seem to balance out the German Gymnastic tradition with that of the Swedish – Lingians, represented in the first International Olympic Committee by Victor Balck, Director of the Royal Gymnastic Central Institute in Stockholm.

¹ Obituary: "A typical Englishman – Dr. W.P. Brookes of Wenlock in Shropshire", in Review of Reviews, New York, 1895.

² "The History of the Liverpool Olympics in 19th Century England"; by Dr. Joachim Ruhl and Annette Keuser, Deutsche Sporthochschule Köln, Cologne, Germany.

³ See illustration: The Olympic Festival at Llandudno: The Feast of Lanterns. This four day Festival comprised athletics, fencing, gymnastics, boxing, wrestling, weightlifting, swimming, and sailing. The Olympic Festival opened with a procession, accompanied by band, and closed with a "Venetian Festival" and dinner-dance. Silver medals were awarded to the winners – bronze ones to runners-up. The Victor Ludorum, a Mr. Henderson from Liverpool and a corn merchant by trade, won one gold medal (the champion), five silver, and five bronzes!



A DATABASE OF OLYMPIANS – the Databasing of Olympic Athletes and Results and Considerations on a Style Sheet of Olympic Names

By Bill Mallon

With the advent of the computer age and the easy accessibility of databases, efforts are now being to create a complete database of the results of all the Olympic Games and of all Olympic athletes. The effort is going forth among several groups, with the JOC also involved. While this database requires numerous decisions to be made concerning the styles of usage for events, nomenclature of codes for sports, disciplines, events, sub-events, and phases, the major problem inherent in it is created by the difficulty in deciding the usage of the athlete's names. In the following I will discuss some of these problems and how they relate to the database, the format of the database, and suggest possible solutions. Much of this material has been reviewed and edited, with many suggestions incorporated from the following ISOH Members: Magne Teigen, David Wallechinsky, Arild Gjerde, Tony Bijkerk, David Foster, and Stan Greenberg.

It may seem trivial to even be concerned about the proper name usage but it most assuredly is not. One prominent American author of self-help books, Norman Vincent Peale, once wrote, "The most important word in any language, is a person's own name." My first Olympic book was co-authored with Ian Buchanan and entitled *Quest for Gold: The Encyclopedia of American Olympians*. At the 1984 track & field Olympic trials, I showed one of the first copies to several Olympians, among the first of which was the javelin thrower, Karin Smith. Her only comment to me, "But you spelled my name wrong." We had used the more common, but in this case, incorrect, Karen.

To me historians who demur that minor fact discrepancies are unimportant, while the arguments they build are important, overlook the fact that if there arguments are built upon an incorrect database of information, the arguments themselves may be invalid.

There are major discrepancies between the way names are listed in Olympic results and databases, and even between varying

results and databases. Having written a number of books, I know that publishers and editor usually create what they term a "style sheet." There may not be a "best" answer to a solution of grammar, spelling, or punctuation. But in a style sheet, the editor or publisher simply makes their best informed decision as to what style to use for some question of grammar or spelling and then always uses it consistently throughout.

Definition of Terms

A number of varying terms are used. We should be consistent in the terms we use, to allow us to define the precise styles more accurately. I will define the terms as I know them. I will use my own name(s) (and my sister's) as examples throughout.

Given Name(s): Also known as Christian name(s), and in the United States as first name and middle name(s). I would recommend using the term "Given Name(s)." My given names are William James.

Surname: Also known as Family name, and in the United States as the last name. I would recommend using the term "Surname." My surname is Mallon.

Hypocoristic. This is defined as the shortened, commonly used form of a name, The hypocoristic of my given name is Bill. David Foster's hypocoristic of his given name is Dave.

Nickname: A special name, also called in the United States, a "pet name," and in some nations as an "eke-name." Examples are Michael "Air" Jordan, Eldrick "Tiger" Woods, and William "Tripp" Schwenk. It is not simply the same as the hypocoristic.

Used Name: This is name by which the person/athlete is commonly known, or the name he/she uses. My used name is Bill Mallon. Many athletes use a name that is different than their full, given names. Few people would recognize William B. Jenner, Susan C. von Saltza, or Frederick C. Lewis, but most of us do know them as Bruce Jenner, Chris von Saltza, and Carl Lewis.

Maiden Name: For females, this is the surname of birth. Thus my sister's maiden name is Mallon. Her birth name was Jessica Mallon.

Married Name: For females, this is the surname used after a marriage. Thus my sister's married name is Spence, or Jessica Spence.

Full Name: This includes the given name, the surname, any nickname or hypocoristic, and maiden and married names. This should only be used in the databases for research purposes, and not in the final listing of the results. My full name is William James "Bill" Mallon. My sister's full name is now Jessica Barbara Mallon Spence.

Particles: Particles are short modifiers to names, often standing for "of," such as "van," "von," "de," "der," and various combinations of the same.

Accent marks/letters: Accent marks are marks that modify the letters of the alphabet and are usually specific to various languages.

This is the same as a diacritical mark. An accented letter is a primary letter modified by an accent or diacritical mark.

Diacritical marks: A mark used to modify a primary letter of the alphabet.

Name order: This refers to whether the name should be listed in the results as Given Name – Surname (most common), or as Surname – Given Name (some Orientals and Hungarians)

Double names: Some athletes have a "double" Surname. These should both be included in the Surname field in the database and included in the Used Name in the results. An example is Thomas Glen-Coats (1908 YAC).

Transliteration: This refers to the "translation" of a name spelled in one alphabet into the primary alphabet in use. The transliteration is language-to-language dependent. That is, transliteration from Greek into French is often different than from Greek into English, even though the alphabets used are the same.

Primary name language and alphabet: This refers to the primary language and alphabet that we will use for the results and database, and for the transliteration language "of destination."

Patronymics and matronymics: This refers to the practice of adding a second Surname to the primary Surname, based on the name of the person/athlete's father or mother.

Titles: This refers to titles given to the person after birth, such as "Count," "Prince," "King," or "Doctor."

Now, given these definitions, I will attempt to explain what I consider to be reasonable rules for use in a style sheet for a database of Olympic athletes.

Names in Results

In the results, the name used should be the athlete's Used Name. Thus, we should use Bill Mallon. In general, the Used Name in the results should only include the athlete's Used Name, without adding modifying initials. Thus the Used Name should be Bruce Jenner and not W. Bruce Jenner, and Chris von Saltza, and not S. Chris von Saltza.

In general, the Used Name will probably use the hypocoristic name, thus Bill Mallon, and not William Mallon. These do not need to be placed in quotes or given other special identification. But if the Used Name includes a nickname, which is not a hypocoristic, it should be included in quotes (or perhaps parentheses). Thus, William "Tripp" Schwenk is the Used Name, including the nickname "Tripp."

There may have to be occasional exceptions to the rule of not including modifying initials, for more precise identification. Some examples of this are as follows: In 1984 at Sarajevo, the United States ice hockey team had two players named David Jensen – they should be identified in the results as David A. Jensen and David H. Jensen. In 1964 the bronze medallist in lightweight boxing was Ronald Allen Harris, while in 1968 the gold medallist in lightweight boxing was Ronald Woodson Harris. They were not the same person, and the listing in the results should be Ronald A. Harris and Ronald W. Harris. And a final example where we

have had to add a special identifier to the athlete's name is the well-known one concerning the two Chinese short-track speed skaters in 1998 at Nagano – both names Yang Yang. They are usually identified as Yang Yang [S] and Yang Yang [A], which I think refers to the month of birth: Yang Yang [S] *September, and Yang Yang [A] *August.

Concerning maiden and married names, we should probably use the name used by the athlete at the time of the Olympics. Thus the result listing for the well-known GDR speed skater should be as follows: 1976 – Andrea Mitscherlich; 1984 – Andrea Schöne-Mitscherlich; 1988 – Andrea Ehrig-Schöne-Mitscherlich. We should not call her Schöne-Mitscherlich in 1976 for ease of identification, because that was not her Used Name in 1976.

Names in Database

As opposed to the results, the database should include all possible names and possibly include fields for each of the possibilities – Given Name, Surname, Nickname, Hypocoristic, Used Name, and for women, Married Name. The Given Name in this case should include all Given Names, thus Kahanamoku should have his Given Name in the database as Duke Paoa Kahinu Makoe Hulikooha. (Incidentally, in this case, "Duke" is not a title, but is his Given Name.) His Used Name is Duke Kahanamoku.

My preference is to have the fields as follows:

Given Name:	William James "Bill"
Surname:	Mallon
Nickname:	Bam-Bam
Used Name:	Bill Mallon
Used Given Name:	Bill
Used Surname:	Mallon

This would include the hypocoristic within the Given Name field. The Used Name could be in a single field, and include the shortened version of the Given Name and Surname, if necessary, as noted above. But David Foster has pointed out correctly, that it is better to have separate fields for the Used Given Name and Used Surname. As he notes, in a database it is easier to have too many fields and combine them, than to try to separate the fields at a later date.

For women, the Surname may have to include separate fields for both the Maiden Name and Married Name. The reason for this is that the rules are different in different nations (see Name order). As you know, in the United States and Great Britain, the female tennis player is known as Chris Evert-Lloyd, but if she was German, her name would have been listed as Chris Lloyd-Evert.

Use of the full name allows genealogical research into athletes and their background. Without a complete name, it is often not easy to find people in genealogical research. For instance, in Denmark trying to find Hans Larsen is almost impossible if that is all you have. But, trying to find Hans Agersnap Larsen is a bit

easier. Ture Widlund and Erich Kamper have told me that it is senseless to do this, because in Sweden and Germany, the middle names are not important. But in many countries they are important, and by including them, the information is available. If somebody does not wish to use them, they may refer to the Used Name, and simply avoid using that information,

I would also only include names used by the athlete(s) in the Olympic Games. Specifically, if a woman marries after the Olympics, I do not think we need a field to list these new additional names, as they do not affect her name listings at the Olympic Games.

Accent Marks and Diacriticals

All accent marks and diacriticals should be included, to the best of our ability to include these in a computerized system. This is currently available in Microsoft Word for Windows 7.0 and Access. The use of Unicode, which defines all letters and alphabets, including diacriticals, is making this easier and will almost certainly become the standard in the future. It appears that Word 2002 will have full Unicode capability and hopefully eliminate future difficulties with accented letters. In addition, Unicode is being made fully compatible for HTML and XML for use on the Internet. We should attempt to use and follow Unicode principles wherever possible. (See Unicode on the Internet for a discussion of this system.)

Particles

Particles should be included in the athlete's Used Name, if he/she includes them as part of the Used Name. They should always be included in the Full Name in the appropriate field(s) in the database.

Problems arise, however, with what is the appropriate field. In the United States, the particle is always included with the Surname, and should be in that field, and the sort order of the database should be based on the diacritical. Thus Susan Christina von Saltza is alphabetized under "von Saltza."

This is not true in the Netherlands, where Tony Bijkerk is insistent that their names should be alphabetized by the "base" of the surname. Thus, Pieter van den Hoogenband should be alphabetized under "Hoogenband" with "van den" included with the Given Name. Arild Gjerde notes that this is also true in Brazil, with all the "de Souza," "dos Santos," etc. Wolf Lyberg does not believe in this system and in his books, alphabetizes under "dos Santos," "von Saltza," and "van den Hoogenband."

There are two possible solutions: 1) always use the same method, probably to include the particle with the Surname in all cases. Thus, alphabetize under "von Saltza," "van den Hoogenband," and "dos Santos," without regard to the nation of origin; or 2) use the method that applies to the nation of origin. Thus, use the system I described above for the United States, but use Tony's system for the Netherlands (and Brazil).

1) has the advantage of consistency and ease of use for somebody using the database who may not understand the differences. It has the disadvantage of not being as precisely correct as 2). 2) has the advantage of being more correct, but it is much more difficult to use. Though it is difficult, I think the second option is more correct.

The next question concerning particles concerns capitalization of them, and this, unfortunately, differs from nation to nation. In the Netherlands, they are never capitalized, thus "Pieter van den Hoogenband." But in Belgium, they are sometimes capitalized, depending on the person's status in earlier times – nobility was not capitalized. Thus, we have Fernand de Montigny but Ivo Van Damme. In Italy, the particle always seems to be capitalized (Roberta Di Donna), while in the German and Slavic-speaking nations, it appears that it is not usually capitalized (Zoltán von Halmay, Cornelius van Oyen).

In the United States and Great Britain, particles are not as common, but they do exist. Unfortunately there is no rule at all and each case must be individualized. Examples of different spellings and styles in the United States include: Chris von Saltza, Amy Van Dyken, Emily deRiel, Albert Vande Weghe.

In general we should likely follow the rules of capitalization of particles (as given above) for each nation, but be aware of individual cases that may not follow those rules.

Titles

Titles should always be included in the database, as they are important identifiers of the person/athlete. They should probably be included as a separate field, again using Foster's theory that it is better to have too many fields than not enough.

Magne Teigen feels that we should not include titles in the result listings, but I disagree for several reasons. 1) Certain athletes (including medallists) are only known by that title – King Konstantine of Greece, Crown Prince Harald of Norway. If Princess Anne had won a medal in equestrian, identifying her as Anne Windsor would be simply wrong, as nobody would even know who that is; 2) Certain athletes (including medallists) are far better known by the title, and by our formula of using the "Used Name" in the results, we should use the title in the "Used Name." The best example is Lord Burghley, or David, Lord Burghley. His given name is David Cecil, but nobody knows him by that name; and finally, 3) Magne used David Northesk, as opposed to David, Earl of Northesk (1928 Skeleton) as an example. But David, Earl of Northesk is better because that is his Used Name. In addition, Northesk may not even be the last name. Using David Burghley is wrong his Given Name and Surname is David Cecil. His Used Name is either David, Lord Burghley, or Lord Burghley.

Patronymics and Matronymics

Patronymics and matronymics should always be included in the database in compiling the Full Name. They should be included in the

Surname field. In the results, the Used Name typically does not include the Patronymic – thus Felix Savón and not Felix Savón Fabre.

Arild has also pointed out differences in Spanish and Portuguese use of the Matronymic. In Spanish, the Matronymic follows the Surname of the father. In Portuguese, it seems to be the other way around.

There are some exceptions where the Matronymic needs to be included in the Used Name in the results. The best current example is the Spanish tennis player Arantxa Sánchez Vicario, who has insisted that she wants her full name, including the Matronymic, to be used. There may also be cases where the Matronymic may be needed to help identify different athletes with the same name, although I do not know of any examples of this currently.

Name Order

This is only a problem for the results. In the database, the order does not matter, because Given Names and Surnames will be contained in separate fields. In the results, the Used Name will contain a portion of the Given Name and possibly a portion of the Surname, or some variation thereof. The order may not always be the same.

Most Western, or Occidental, nations use the order "Given Name-Surname." Thus, I am Bill Mallon. But certain Oriental nations use "Surname-Given Name," without a comma in the middle. Thus, in China I would be Mallon Bill. This system is also used in Hungary thus Kovács István, and not István Kovács, but apparently they are not as strict about this and the Occidental order is acceptable.

The Oriental order system is used by China, Korea (both), and Chinese Taipei, for certain, and technically Japan, although they seem not follow the rule closely and actually tend to use the Occidental style at the Olympics. What I am uncertain of is whether or not this system is used for other Pacific Rim nations with names similar in "style" to the Oriental nations. I am not positive of the name order used for the following nations, although I suspect it follows the Oriental rule: Hong Kong, Vietnam, Malaysia, Singapore, and Indonesia.

I think we should follow the order used by the nation in question. Thus most nations would have the Used Name in the order "Given Name-Surname," but China, Korea, and Chinese Taipei (and perhaps several of the others) would use the Oriental system. I would recommend that we continue to use the Western/Occidental system for Hungary. (See also Name Style [below])

Transliteration

The primary alphabet we have to use is obviously the Latin alphabet. The primary language should likely be the English language, because of its universality. The IOC may demur and prefer French, however. But using my style sheet, transliteration should always be done from the original alphabet into the English language and Latin alphabet. It is not a simple matter to perform these transliterations and they will vary depending on the language chosen.

Thus, the Russian name spelled Iouri in transliteration into French, or Jurij transliterated into German, is often seen spelled Yuri, Yury, or Yuriy on transliteration into English. It depends somewhat on the Transliteration Table one uses, and there are several.

The *Encyclopædia Britannica* has a transliteration table for most Cyrillic and Greek alphabets. It is a well-respected source and it gives the transliteration tables into English and the Latin alphabet and I think makes the best choice as it is also more easily accessible than many of the other tables.

The primary languages in question here are Russian and Greek. For the other languages in the former Soviet republics, the problem becomes much more difficult. The Ukraine has a very different language, although it still uses the Cyrillic alphabet. Care must be taken to do these transliterations using the tables for Ukrainian and not Russian. Other languages, specifically Armenian and Georgian, actually use a different alphabet, and I have no good answer, except to use the original spellings if they can be found, and transliterate according to a well-accepted table.

For the Baltic nations the problems is very difficult. These nations do not use the Cyrillic alphabet and many of the Baltic athletes who competed for the former Soviet Union had their names transliterated by Soviet sports authorities into Cyrillic and Russian. When Western journalists transliterated the name back into Latin, it often came out spelled differently from the original.

Bulgarian, Serbian, and Macedonian are similar to Russian in that they use the Cyrillic alphabet, but they are different. There is a separate transliteration table in the *Encyclopædia Britannica* for Bulgarian and Serbian.

What about other languages and alphabets? Here we have even more problems. I refer to Arabic, Farsi, Chinese, Korean, Japanese, Thai, Turkish, and similar difficult alphabets (to us Westerners). I think most of us would even have trouble reading these alphabets to do the transliterations and I think we may have to rely on local experts. For example, Esmail Kadkhosadeh has helped me with Iranian names, and Cuneyt Koryurek has provided all the Turkish name spellings. We will need to get help from such experts and rely on it.

In Chinese, the current transliteration system must be used, which is Pinyin, as opposed to the older Wade-Giles system. Thus, the currently accepted spelling is Mao Zedong, and not the older Mao Tse-Tung. In Korean, the system used is the McCune-Reischauer system, which mandates the name order and style discussed below.

Name Style

For several of these transliterated names, the style of how the name is presented is important. Primary examples discussed are Chinese, Korean, and Arabic names.

For Chinese, the "standard" style seems to be (at least in Pinyin) to use the name order "Surname-Given Name" as discussed above. But the given name is now usually spelled as one word, with no hyphen. Thus, Mao Zedong and not Mao Ze-Dong. This appears to be preferred for Chinese names, but it is more variable in Chinese Taipei.

For Korean, in the McCune-Reischauer system, the style includes the Oriental name order, but the Given Name is split into two and usually hyphenated, and usually the second part of it is capitalized. Thus Kim Soo-Nyung is probably best, as opposed to Kim Soonyung, or Kim Soo-nyung. I prefer this style for Korean names and probably also for DPR Korean (North) names, although it is more variable in that nation.

For Arabic, the opening Surname particle is very often "Al" or "El." The spelling of this is very variable. I have seen Mohammed Alkhaidi, Mohammed Al Khaidi, Mohammed al Khaidi, Mohammed al-Khaidi, and Mohammed Al-Khaidi. I do not know of a precise style or system for Arabic, but I would recommend that we be consistent. I would recommend using "Al-K..." or "El-K..." in all cases, thus the correct style above would be Mohammed Al-Khaidi. It appears that most Arabic nations use "Al-..." but Syria seems to use "El-..." primarily.

For married/maiden names, I have described these above. The order varies by nation, but I would tend to use the order of the nation in question, because it tends to be the athlete's Used Name. Thus Chris Evert-Lloyd was her Used Name, when she was married to John Lloyd. But Heike Daute became Heike Drechsler-Daute after her marriage, because the German system is different.

The Tchaikovsky Question

Linguists and scholars who deal with transliteration questions have a rule they call "The Tchaikovsky Rule," referring to the Russian composer, Pyotr Ilyich Tchaikovsky. There is no known transliteration system that would make Tchaikovsky's name be spelled in this manner. The probable best transliteration into English, using the *Encyclopædia Britannica* tables, is Tschaykovsky, but nobody ever uses that.

The Tchaikovsky Rule is that "Once a person/athlete's transliterated name becomes established, by whatever method of transliteration, it becomes the Used Name." By this system, we should continue to use the Used Name, even if transliterations seem to be wrong.

Two good examples are the Moroccan miler Hicham El Guerrouj, and the Russian figure skater Ekaterina Gordeeva. By the rules I have outlined above, these would/should be Hicham El-Guerrouj and Yekaterina Gordeyeva. But they are fully known by these Used Names and we should probably continue to use their commonly seen names, following The Tchaikovsky Rule.

Summary

There are many variables to consider in the spelling of the names of athletes from the many nations of the world. The naming rules are difficult and vary from nation to nation, from language to language, and from alphabet to alphabet. Ideally, we would always spell the name as the athlete does, but it is not always possible to contact the athlete and know this precisely. When not possible, in general, we should try to use the rules and spelling of the nation of the alphabet involved in all cases.