

# Indexing the Names of Authors from Spanish- and Portuguese-Speaking Countries

## Bill Black

Journals in the United States and other English-speaking countries—as well as journals written either entirely or partly in English in other nations—are receiving growing numbers of papers from authors whose first language is not English. For example, Latin America's world share of scientific literature indexed by ISI more than doubled from 1981 to 2000, and it continues to climb.<sup>1</sup> In addition to the challenge of helping authors from around the world to express their ideas in English, editors have the responsibility to make certain to present and index their names correctly. Besides being a courtesy to the authors, it can help to ensure that other researchers and bibliographic database services that cite the articles give the credit due to the authors and to the editors' journals and that other researchers who see the piece cited can easily locate abstracts of the articles or their full text. Knowing a little more about names from countries where English is not the first language can also help in reviewing the reference items in the articles.

Understanding and correctly writing names from another language can be difficult, as was discussed in the case of Chinese in an earlier *Science Editor* article.<sup>2</sup> There are similar challenges with the names of authors from Spanish- or Portuguese-speaking countries. Both those languages traditionally use naming patterns that are different from the ones in English-speaking nations. In addition, the rules in Spanish differ from those in Portuguese in ways that often make Spanish names more difficult than Portuguese ones to index correctly.

In this article, I present the "classical" rules about Spanish- and Portuguese-lan-

guage names; discuss how authors write their names, as shown by submissions to our own journal over the last several years; provide guidance on how to "read" the names so that you can index them correctly; and offer some guidelines that you might want to suggest to authors and some changes you may want to make in your journal to increase the chances that others will cite your authors and your articles correctly.

## The Classical Rules

The style manual that we use for our journal provides basic, "classical" rules for Spanish and Portuguese names.<sup>3</sup> Names of Spanish origin, it says, often use two surnames, the first being the father's and the second the mother's. Almost without exception, such "compound" surnames should be alphabetized and indexed according to the father's name. When a woman marries, she drops her mother's name but keeps her father's name, to which she appends the preposition *de* ("of") followed by her husband's father's surname. For example, if María Esquivel López marries Juan Sánchez Mendoza, she drops her mother's surname (López), keeps her father's surname (Esquivel), and adds "de Sánchez," becoming María Esquivel de Sánchez. If María and Juan have two children, a daughter named Susana and a son named Jorge, their full names would be Susana Sánchez Esquivel (that is, her given name, then her father's surname and her mother's surname) and Jorge Sánchez Esquivel (that is, his given name, then his father's surname and his mother's surname).

There can be some variations in how the parts of the compound surname are presented in Spanish. Some women omit the *de* when adding the husband's surname, such as using María Esquivel Sánchez instead of María Esquivel de Sánchez. Both men and women may use the conjunction *y* ("and") to join the paternal and maternal names,

writing Juan Sánchez y Mendoza instead of Juan Sánchez Mendoza. Some persons use a hyphen to join the two surnames, such as Juan Sánchez-Mendoza or María Esquivel-Sánchez. Some men and women use just an initial for the maternal surname, such as Juan Sánchez M.

Portuguese also uses compound surnames, but with the order reversed from that in Spanish, that is, with the mother's surname coming first and the father's surname second. Women vary in how they write their name after becoming married; a common practice is to drop the maternal surname and keep the paternal surname and add the husband's paternal surname. In all those cases, the indexing of the Portuguese names would use the last word in the author's name, much as it is in English. (Later I'll consider some specific exceptions or confusing situations, such as the Portuguese equivalents for "Junior" [*Filho*] or "III" [*Neto*].)

Both Spanish and Portuguese also often use compound given names, such as María Cristina or José Fernando in Spanish or Maria Tereza or Lúcia Helena in Portuguese. Compound given names can also include prepositions and articles, such as María de Jesús or María de las Mercedes in Spanish and Maria das Dores or Isabel da Conceição in Portuguese. Long names in Portuguese are no problem for indexing, because they are still indexed by the paternal surname, which will be the last word in the name, as it is in English. However, for persons unfamiliar with Spanish names, the use of compound given names can make it difficult to decide whether a word that falls in the middle of a name is the second part of the given name or is the paternal surname. Later I present some guidelines to help in making those distinctions.

## Our Sample vs the Classical Rules

Are authors following the classical rules? To assess that, we looked at the English-

BILL BLACK is the editor of English-language materials for *Revista Panamericana de Salud Pública*/Pan American Journal of Public Health).

*Indexing continued*

language articles submitted to our journal, *Revista Panamericana de Salud Pública/Pan American Journal of Public Health*, over the 3-year period January 2000 through December 2002. Our journal carries articles in English, Spanish, and Portuguese; each article appears only once and in only one of the three languages. The large majority of authors are from the western hemisphere; a very small number are from other parts of the world, such as Spain. We encourage authors to write in their first language—and thus their strongest one—but they may choose to write in another one. For example, an author from Brazil might decide to write in English or Spanish instead of his or her native Portuguese.

We included all authors of English-language papers who listed affiliations with institutions in Spanish- or Portuguese-speaking countries. To that group we added authors with Spanish or Portuguese names who were affiliated with international organizations based in English-speaking countries and for which Spanish or Portuguese is one of the official languages. (That was most commonly true for authors affiliated with our journal's parent organization, the Pan American Health Organization.)

We excluded authors with Spanish or Portuguese names who were affiliated with other types of institutions in English-speaking countries, such as universities. Such authors may have been born in English-speaking countries, or they could have been working for some time in English-oriented institutions. In either case, we thought those persons would be more likely to be following the standard English-language rules for names, so their names could be indexed in the same way as standard English names.

We chose to use our journal to compile the sample of names because we believed that the authors who write in English for our journal would be representative of the wide array of Spanish- or Portuguese-language authors writing for English-language journals in the United States and elsewhere. Some of our authors have received

advanced degrees in English-speaking countries, write well in English, and may be familiar with the differences between English names and those in Spanish or Portuguese. Others have never lived outside their home countries, have limited English writing skills and so pay to have professionals translate their articles into English, and may not know the English naming rules. Furthermore, although they may have published extensively in Spanish or Portuguese in their home countries or elsewhere, this may be the first time that they are writing an English-language article for an international journal.

Are all our authors following the "classical" Spanish and Portuguese rules for names? In a word, no. In fact, much to the relief of editors of English-language journals, most of the authors whose names we surveyed are using forms of their names that will present few problems for English-language editors. For example, of the 90 Spanish names checked, 41 (45.6%) consisted of just a single given name and a single surname (for example, Susana Vázquez). Another 15 (16.7%) consisted of a single given name, an initial, and one surname (for example, María G Guzmán). (A check of Spanish names from Spanish-language articles that we published in the year 2000 showed similar percentages, of 54.4% and 10.7%, respectively.) Another six (6.7%) consisted of one given name and a hyphenated compound surname (for example, Héctor Díaz-Torres). All three of those categories could be easily indexed by following the customary practices for English-language names. Where there will most likely be problems in Spanish are with names that use three or more words without hyphens or that have conjunctions, prepositions, or articles. (Because relatively few Portuguese names will present difficulties, I do not provide data on the various forms that we found in our English-language articles; however, some of the most common problems are discussed below.)

### **Distinguishing Spanish from Portuguese**

How do you and your journal deal with

this variety of practices when authors do not use a form that is close to the standard practices for English-language names? Two steps are essential: 1) distinguishing Spanish from Portuguese and 2) learning to read the "elements" of a Spanish or Portuguese name, such as distinguishing between given names and surnames and knowing how to handle prepositions and articles. (There's also a guideline, presented at the end of this article, that may supersede the above.)

Of the major Romance languages, Spanish and Portuguese may be the two that are most similar. In distinguishing between Spanish and Portuguese, probably the best indicator is the author's country. Spain and Portugal colonized much of the New World and brought their respective languages with them. With the exception of Brazil, where Portuguese is spoken, all the larger countries of Central America and South America use Spanish. In the Caribbean, the most common languages are English, Spanish, and French; it depends on which country colonized the island. If you are not sure of the official language of a country, a quick look at an almanac or a search on the Internet can answer the question.

Another way to distinguish between Spanish and Portuguese is to look at the use of accents and other diacritical marks. For example, Spanish uses "ñ" (n with tilde, as in the surname Núñez), but Portuguese does not. Portuguese uses the "ç" cedilla form (as in the surname Gonçalves), but Spanish does not.

Both languages use accent marks over vowels, but Spanish uses only one form and Portuguese uses several. Both languages use "acute" accent marks, such as the "é" in José, a given name in Spanish and Portuguese. Portuguese also uses the circumflex form, such as the "ô" in Antônio, and the tilde over a vowel to indicate its nasalized form, such as João, the Portuguese form for John.

A complete list of the diacritics used in Spanish and Portuguese is given in *The Chicago Manual of Style* (see the chapter

titled “Foreign Languages in Type”). However, it is not necessary to memorize all of them. With a little experience, one will probably be able to look quickly at a name, see the spelling and diacritical marks used, and have a good idea of whether it is Spanish or Portuguese. That is much the same way that a person who has heard some French and some German can distinguish between the two by their sound patterns without being able to understand the words themselves. For example, one can quickly learn to distinguish between the Portuguese João Gonçalves and Antônio Nunes and the Spanish Juan González and Antonio Núñez.

### Identifying the Elements of a Name in Spanish

With a Spanish name that has more than one given name or more than one surname, how do you recognize which is which? For example, how do you “read” the name of María Isabel Gutiérrez? Is Isabel part of the given name, or is it the paternal surname that should be used for indexing? What about Fernando Ruiz Meir? Is Ruiz part of the given name, or it is the paternal surname that should be used for indexing?

As noted earlier, if the author has hyphenated the two surnames, that simplifies the task. For example, in the case of Juan Sánchez-Mendoza, the author would be indexed by the compound surname Sánchez-Mendoza (the abbreviated form would be Sánchez-Mendoza J). (The recommendations for indexing are based on the rules for names given in *Anglo-American Cataloguing Rules*, 2nd Revision, AACR2, a text that most libraries in the United States and other English-speaking countries apply in indexing materials, and on standard indexing rules in Spanish- and Portuguese-speaking countries. The abbreviated forms suggested here are based on the full names and the appropriate abbreviations according to the “Vancouver style” for references in manuscripts submitted to biomedical journals.<sup>4</sup>)

A variety of patterns can help distinguish given names from surnames. Some Spanish given names are written the same as in English, such as Linda, Patricia, Irene,

David, and Daniel. Others are similar but use an accent mark, such as Martín, or have a small difference in spelling, such as Eduardo and Sonia. Men’s given names in Spanish can often be identified by their –o ending, such as Fernando, Rodrigo, and Gonzalo. Women’s given names frequently end in –a, such as María, Cristina, and Anita. Another way to identify a given name is to use a guide like the ones that prospective parents use to choose a baby’s name. Several English-language guides to Spanish-language names have been published, and a number of similar Internet sites can be found by doing a Google search for “Spanish baby names”. One useful rule for identifying surnames is that many of them are formed by adding an –ez to a given name, such as Rodríguez, from Rodrigo and González from Gonzalo. For the first two examples given above, María Isabel Gutiérrez should be indexed by the single surname Gutiérrez (the abbreviated form would be Gutiérrez MI), and Fernando Ruiz Meir should be indexed by the compound surname of Ruiz Meir (the abbreviated form would be Ruiz Meir F).

Spanish surnames that use conjunctions, prepositions, and articles can be confusing. Among the most common of those forms are the conjunction *y* (“and”); the preposition *de* (“of”); *del*, *de la*, *de los*, and *de las*, which all combine a preposition and an article; and *Los* and *Las*, which are articles (and which some authors will write as *los* or *las*).

The preposition *de* and the related preposition-article combinations (*del*, *de la*, *de los*, and *de las*) can appear with the paternal and/or maternal surname. As in other Spanish names, the author’s name should be indexed by the paternal surname; for example, José del Pozo García should be indexed by Pozo (the abbreviated form would be Pozo García J del). The abbreviated form keeps the full form of the preposition or preposition-article combination but puts it at the end of the name. Similarly, for a maternal surname using *de* or one of the combined forms, the author’s name should still be indexed by the paternal name; for example, Juan González del Moral would be indexed by

González (the abbreviated form would be González del Moral J).

When used without a preposition, the articles *Los* and *Las* should remain with the surname; for example, the author Manuel Antonio Las Heras should be indexed by Las Heras (the abbreviated form would be Las Heras MA).

### Identifying the Elements of a Name in Portuguese

In general, Portuguese names present fewer difficulties than Spanish names. Many authors use a form that is very similar to English: one given name, one surname, and perhaps one or two initials between them. Even if the author uses a compound given name and a compound surname, the last word in the name will generally be the one used for indexing, as is normally the case in English. For example, Neusa Maria Costa Alexandre would be indexed by the paternal surname Alexandre (the abbreviated form would be Alexandre NMC).

As in Spanish, there can be difficulties with prepositions and articles or combinations of them when they form part of the paternal surname. Those words include the prepositions *de*, *da*, and *do* and the preposition-article combination *dos*. For example, should the name Boaventura Antônio dos Santos be indexed by Santos or dos Santos?

If the paternal surname uses any of the *de*, *da*, *do*, or *dos* forms, the name is indexed by the word that follows. For example, Boaventura Antônio dos Santos should be indexed by Santos (the abbreviated form would be Santos BA dos).

With surnames and the conjunction *e*, usage varies widely. Sometimes *e* joins two separate surnames, and at other times the surname + *e* + surname formation has become “fixed” as a single surname. The same author may use the conjunction at some times and not at others. With this *e* conjunction, it is best to ask the author what he or she prefers.

Another difficulty in Portuguese is with the forms that follow the surname and are equivalent to “Junior” (*Filho*) and “III” (*Neto*). They should be kept as part of the surname, not treated as the final, paternal

*Indexing continued*

name itself. Some authors make this clearer by putting a hyphen between the surname and the *Filho* or *Neto* form, as in Djalma C. Moreira-Filho, which would be indexed by Moreira-Filho (the abbreviated form would be Moreira-Filho DC).

Other authors, such as Alberto Pellegrini Filho and Antônio Ribeiro de Castro Neto, do not use a hyphen. Those two authors should be indexed by Pellegrini Filho and Castro Neto (the abbreviated forms would be Pellegrini Filho A and Castro Neto AR de).

**Guidelines for Your Journal**

Are there any general guidelines that you can offer to authors that will make your life as a journal editor easier? For Spanish names, several helpful ideas were suggested in a recent article by Ruiz-Pérez and others.<sup>5</sup> Authors should be encouraged to use the same name throughout their publishing career. It is preferable to use a form that is close to the normal English patterns: one or two given names and one surname. Prepositions, articles, and conjunctions should be avoided. When two surnames are used, they should be joined with a hyphen. Another suggestion for Spanish authors, not listed by Ruiz-Pérez and others, is to discourage authors from using a single initial at the end of their name in place of the full maternal surname (such as Carlos Suárez N); it would be better to use either just the paternal surname or both surnames joined by a hyphen.

Several of those suggestions are also applicable for Portuguese names. For example, authors should be encouraged to use just one form of the name for professional work. Authors using *Filho* or *Neto* can join it to their surname with a hyphen.

Depending on how many manuscripts you receive from authors whose first lan-

guage is not English, you may want to include some of these recommendations in your journal's guidelines. Or, it may be easier just to raise these kinds of questions along with the others that you send to authors concerning their manuscripts.

What if you have tried to follow these guidelines and are still not sure what to do with a particular name? Then follow our special, third rule: Ask the author what form he or she wants. Besides making certain that your journal will index the name properly, both you and the author are likely to benefit from this modest cultural exchange.

**Getting Others to Follow Your Example**

That your journal is indexing Spanish- or Portuguese-language names appropriately does not mean that the correct forms will be reproduced by others who cite or index your articles. For example, a researcher could read the full text of an article, see the authors' full names, and then cite their names incorrectly if he or she doesn't know the correct rules for two nonhyphenated surnames in Spanish. The same can happen if a Portuguese-language author hasn't used a hyphen to connect *Filho* or *Neto* to the surname. Major database and indexing services frequently make mistakes with Spanish names, as Ruiz-Pérez and others point out,<sup>5</sup> and a quick search can show a number of questionable *Filho* and *Neto* Portuguese "surnames" in such databases as PubMed.

You might consider a change that we have decided to make for our journal. In addition to presenting a list with the authors' full names, we will add to each of our articles a suggested Vancouver-style citation form, with the authors' names

given in their correct abbreviated forms. We will include a draft version of that citation in the edited version of the article that we send to the authors for their review and approval, so we can be certain to use the form of names that they want.

Learning how to index Spanish- and Portuguese-language names properly and then explicitly telling others how they should cite your articles should increase the chances that both your authors and your journal receive the credit that they are due. 

**Acknowledgments:** Thanks go to Ana Leone and to María Luisa Clark, who reviewed the manuscript. Leone handles the indexing for *Revista Panamericana de Salud Pública/Pan American Journal of Public Health*, and Clark is the managing editor of the journal.

**References**

1. ISI. Latin America: a growing presence. *Science Watch* 2001;12(5). Available at [www.sciencewatch.com/sept-oct2001/index.html](http://www.sciencewatch.com/sept-oct2001/index.html). Accessed 20 March 2003.
2. Sun XL, Zhou J. English versions of Chinese authors' names in biomedical journals: observations and recommendations. *Sci Ed* 2002;25(1):3-4.
3. Pan American Health Organization. PAHO style manual. Washington, DC: Pan American Health Organization; 1994.
4. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. Available at [www.icmje.org](http://www.icmje.org). Accessed 20 March 2003.
5. Ruiz-Pérez R, Delgado López-Cózar E, Jiménez-Contreras E. Spanish personal name variations in national and international biomedical databases: implications for information retrieval and bibliometric studies. *J Med Libr Assoc* 2002;90(4):411-30.